United Nations ECE/CES/2015/37

Economic and Social Council

Distr.: General
22 May 2015

English only

Economic Commission for Europe

Conference of European Statisticians

Sixty-third plenary session
Geneva, 15-17 June 2015
Item 4 of the provisional agenda
Modernization of statistical production and services and managing for efficiency

Business driven improvements in Statistics Norway

Note by Statistics Norway

Summary

The paper describes the new strategy for the development of statistical methodology in Statistics Norway. The strategy emphasizes standardization work and the need for close cooperation between statisticians, methodological experts and information technology specialists. The paper will highlight the measures taken to enhance strategic governance, enterprise architecture and agile capabilities in statistical production.

The paper is presented for discussion to the first session of the Conference of European Statisticians' seminar "Modernisation of statistical production and services and managing for efficiency".

GE.15-10385 (E)
1510385

Please recycle

I. Introduction

1. Like many other national statistical institutes, Statistics Norway has to meet new user demands, competition and needs for improved efficiency. At the same time new data and technology can provide new possibilities. The strategies and plans of Statistics Norway address meeting these challenges. They aim to streamline the production of statistics, emphasizing the core values for official statistics as they are reflected in the United Nations Fundamental Principles of Official Statistics and the European Statistics Code of Practice.

2. Comprehensive programs to improve the business agility through quality management, Lean methodology (see Section III) and streamlining statistical production are in progress. Effective tools should be available to support all parts of work processes. A challenge is that the speed of development and change is not yet satisfactory. Even though the major development programs are regarded as top priority, they also compete for scarce resources and skills, and do not always create sufficient synergies across programs. There is a need to coordinate the major developments better, and to be able to prioritize projects across the different programs.

3. An improved holistic view of the business needs, and a better understanding of the business processes and interactions between business functions are necessary to improve the speed of change. This has led to the building of capacity in enterprise architecture and a need for more speedy development of production systems.

4. New business competencies and governance strategies and processes must be developed accordingly. This could lead to changes in the way we organize and run the business of official statistics. Having centralized the information technology (IT) function, the responsibility for strategic business developments is still to some extent decentralized. No specific business function has been given the cross sectional responsibility to coordinate and lead the strategic business process developments.

5. These issues are the main focus of this paper. The paper will highlight the measures taken to enhance strategic governance, enterprise architecture and agile capabilities. It begins with a brief overview of the status of improvement work with regard to quality, Lean programme and streamlining statistical production.

II. Work on quality

6. Statistics Norway has focused on the importance of quality management during the last decades. A systematic approach based on common principles of quality management has been adopted. The work has been supported by international initiatives, in particular the European Statistics Code of Practice (CoP) for the production and dissemination of statistics. A second round of peer reviews assessing compliance with CoP has just been completed.

7. Tools and procedures applied for quality assurance in Statistics Norway include quality reports and indicators, user surveys, self-assessments and quality reviews[1]. Where relevant, tools can be linked to different stages in the business process model. Together with a quality framework, a business process model and an organisation for coordinating work on quality constitute a necessary infrastructure for systematic work on quality in a statistical institution.

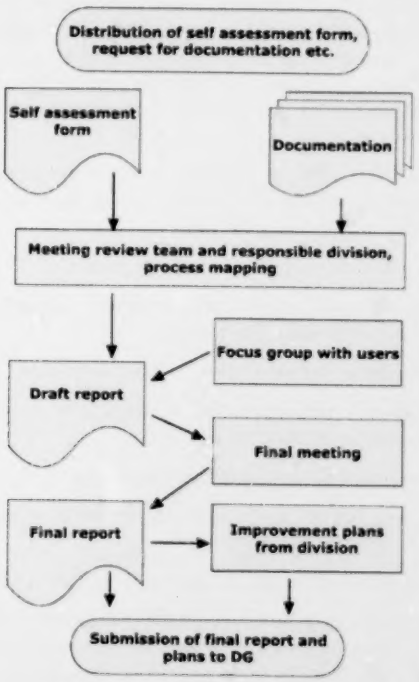
[1] Quality Assessment and Improvement Methods in Statistics – What works Sæbø, H.V., Statistics and Economy Journal, Vol. 94 (4) 2014, Prague 2014

A. Quality reviews

8. Quality management is important not only to assure quality when changing business processes, but also to identify potential areas for change. Quality reviews are an efficient tool to achieve this.

9. Statistics Norway started work with systematic internal quality reviews or audits of selected statistics in 2011[2] (see also footnote 1). CoP has guided the reviews. The reviewing process is illustrated in figure 1. It is performed very much like the European peer reviews, with the exception that specific statistics or subject matter areas are reviewed and not the institution as such.

Figure 1.
Setup of quality reviews in Statistics Norway


Distribution of self assessment form, request for documentation etc.
Self assessment form
Documentation
Meeting review team and responsible division, process mapping
Focus group with users
Draft report
Final meeting
Final report
Improvement plans from division
Submission of final report and plans to DG

10. In the period 2011 – 2013, 21 different statistics or areas of statistics have been reviewed; at least one in each division producing statistics. Together these statistics represent almost 30 percent of the working hours used for statistics production in Statistics Norway. In 2014, no reviews were implemented, partly due to an escalation of the Lean project in the organisation. One of the recommendations from the European peer review is to implement regular quality reviews, and these have been taken up again in 2015. They are adapted to the Lean programme to prevent duplication and to supplement this.

[2] Quality reviews in Statistics Norway – Setup and experiences Sæbø, H.V and Byfuglien, J., Paper for the Nordic Statistical Meeting, Bergen, August 2013.

11. The internal quality reviews have been based on three elements: Self-assessments on compliance with the principles and indicators in the CoP and other documentation, process mapping using Lean techniques and focus groups to evaluate user needs. A team of 4 persons has conducted the review. The team members have a background in quality management, statistics production, dissemination and survey methodology. One of the members is a methodologist. The team has been assisted by experts in conducting focus groups.

12. The reviews were "audit –like" even if they were carried out by an internal team. This implies focus on evidence. Findings are presented objectively in a report that is the sole responsibility of the team. The reports follow a standard structure, also including a consideration of strengths and weaknesses. There is no ranking, but each report ends up with a set recommendations based on the findings. The division responsible for the statistics reviewed can correct factual errors, but will make a separate action list on the basis of the recommendations. If they disagree with some of the findings/recommendations they can express this here. Reports and action plans have been sent to the Director General and are followed up annually. They are published on the Statistics Norway Intranet. The different steps in the reviews and experiences are described in more detail in the Quality Reviews in Statistics Norway (see footnote 2).

13. The reviews have resulted in more than 170 proposals for improvements. Many of the improvement points concern several statistics, and there is a reason to believe that they are valid generally in Statistics Norway.

14. The most important improvement points concern the need for:

- More focus on user needs and the relevance of statistics, in particular users want to see the statistics in a broader context;
- Better dissemination with more visualisation (graphs, maps, etc.);
- Improved documentation, in particular a need to update "About the statistics" which provides metadata for the users on www.ssb.no;
- Improved production processes – use of standardised tools;
- Increased understanding for and use of statistical methods, in particular in editing which requires relatively large resources;
- Increased knowledge of formalities (such as the basis for data collection).

15. Most of the improvement actions based on recommendations from the quality reviews performed before 2014 have now been completed. However, what remain are actions that depend on other parts of the organization than the subject matter divisions. For instance recommendations from the reviews emphasize the use of standardized tools, but the lack of speed in developing the common functionality that is needed can lead to costly changes in legacy systems. Furthermore, the resources available to continue the general quality work are scarce. These are examples of actions that require coordination, linked to the other major developments and issues addressed in this paper. For instance the Lean project would benefit from a close cooperation with other quality work, and especially with other staff experts on production processes, statistical methods, etc.

III. Lean programme

16. The purpose of the Lean programme is to create a culture for continuous improvement in Statistics Norway, with the main focus on improved timeliness and reduced use of resources. The Lean initiative follows the lines of earlier work on

Total Quality Management (TQM) and other past efforts to enhance quality. A comprehensive programme to implement Lean methodology in Statistics Norway started in late 2011.

17. The core of the Lean methodology, as introduced in Statistics Norway, is to systematically assess work processes and operational control. Even though the goals are set by management, all staff is invited to participate in improving work processes. This gives them the possibility to develop their own roles and positions and influence the progress.

18. Tools for better operational management and control are central. All departments are involved, and mapping of processes to identify bottlenecks is carried out. To facilitate this, a number of internal Lean guides have been trained. Some 50 projects have been carried out by now, in "waves" that comprise a few departments at the same time. The work is first concentrated on describing and proposing improvement of processes within each department. The objective is to train the staff in Lean thinking and techniques in order to establish a Lean culture. The next phase will go deeper into processes that involve several departments, where we will look at the possibilities to improve core and cross-cutting processes. These are more costly and challenging, but normally have a greater potential for improvement and benefits.

19. Again, this addresses the need to coordinate different initiatives. Within the first phase of Lean, recommendations from the quality reviews have been used as criteria to choose areas or work processes for Lean improvement projects.

20. When we start improving cross-cutting processes in the next phase of Lean, all major development and standardization projects such as those mentioned in the next chapter have to be taken into account, also the work on establishing enterprise architecture. The business process model will be a useful tool in this context.

21. The Lean programme has an internal project manager, but external consultants are currently quite heavily involved in progressing Lean in the first phase. This is useful to ensure a sound general methodology that works in different organisations, and sufficient force in the implementation of the program. However, compliance with frameworks and practices that are specific for national statistical institutions have to be safeguarded by our own experts.

IV. Streamlining statistical production

22. Early work on the development of standard IT solutions in Statistics Norway started around the year 2000. However, more coordinated developments to support the complete model of working processes in the production of statistics within Statistics Norway started in 2008. From the beginning the work was steered and coordinated within the "FOSS" project[3]. One of the first tasks of that project was to develop a complete business process model. Statistics Norway chose a stepwise method of carrying out development work, where we built partly upon solutions that already existed, or on developments already commenced which could be developed further into general and shared solutions. Nevertheless, most of the developments were upon completely new solutions.

[3] Standardisation for improvements in Statistics Norway (Gløersen, Rune and Sæbø, H.V.), Paper for the seminar on Modernization of Statistics Production, Stockholm 2009

23. Statistics Norway has chosen a strategy to develop four integrated production platforms, which to a large extent cover the main phases Collect, Process, Analyse and Disseminate. We have developed a central data reception system that is well integrated with the Norwegian Government portal for web questionnaires, through which Statistics Norway will collect all web forms from businesses. Furthermore, we have established a central system for data editing and estimation, which currently is used by approximately half of the sample surveys carried out. The system is used both for business and household surveys, and has just recently been enhanced to cover also register based statistics. We have also established a first version of a data warehouse solution for analysing and producing tables for register based statistics on labour market and earnings. This solution will be developed further towards an enterprise warehouse for Statistics Norway. Finally, the dissemination platform has been completely modernized, and we are currently improving the Statistical databank, which serves as the master data store for all data published on the web site. In addition, a complete and integrated metadata system is needed. This is only partly implemented.

24. The need to ensure coherence within the new production system and to provide for the possibility to break down developments into smaller modules, has led to increased efforts to define the target systems architecture. Developments should be carried out according to this target architecture, to ensure the integration needed within and between the main production platforms. The need for further work in this area is described later in this document.

25. To underpin the need to speed up change, Statistics Norway is currently about half-way through the developments needed to provide for a complete new platform for the production of statistics. Concerning data collection, some collection instruments, for instance Blaise, are not well integrated into the central data reception system. Basically this is a problem regarding the need to share metadata for further processing. The most complete platform is within dissemination. However, the functional and technical backlog in this area is also rather comprehensive. Within data editing, there is a need to implement more sophisticated editing functionality, especially on statistical edits, macro editing and automated corrections. The master metadata systems will be upgraded, and a comprehensive work to ensure complete metadata integration throughout the production process and life cycle of data is still just beginning.

26. Shared solutions are dependent upon standardisation at different levels. Development of shared solutions requires a common terminology for concepts, which again is a precondition for standardisation and reuse of variables, classifications, data descriptions and the use of technical standards for representation of data. Statistics Norway has therefore invested a lot of work in the preconditions for developing and storing shared metadata. A master plan for improving information management at corporate level is under way.

27. The process of changing the production systems from supporting the production of individual statistical products, to support the functionality needed by common production processes, has led to a fundamental change in the skills needed for planning and development. To ensure coherence at the enterprise level and to provide for the focus needed to support the strategic direction set out by top management, the capabilities within strategic governance will be strengthened.

V. Strategic governance

28. The comprehensive development programs mentioned in this paper, emphasize the need to improve the capability within strategic leadership and

governance. Measures to establish the discipline and competence of design by architecture, and the need to run project development in a controlled environment have been taken. To some extent, the focus on the general architecture, the standardisation of processes and tools, and the need for a strict regime of project and portfolio management could be regarded as unnecessary bureaucracy and an impediment to the agility needed for rapid change. Our view is that these are part of the solution and not part of the problem.

29. Streamlining statistics production by standardising the tools and methods according to general work processes can hardly succeed, unless

- The user needs are thoroughly analysed;
- The development work is driven towards a target architecture;
- Sufficient architecture capabilities and governance are in place;
- Projects are strategically governed in a controlled environment of portfolio management;
- A framework for programme governance is in place to ensure coherence across projects and to optimize the total benefits.

30. Developments should be guided towards well-defined and communicated target architecture. This will help to decompose comprehensive development projects into smaller parts, providing for stepwise changes, which significantly reduce the risk of failure. The target architecture is the basis for ensuring the necessary coherence and integration. Continuous deployment of new features improves the business ability to adapt to change, and forces the implementation of change management to become a business wide competence.

31. As the business changes over time, so will the architecture. Defining and building up the enterprise architecture is in itself a stepwise process. It takes time, and all parts are not necessary from the beginning. However, the complexity of the business and the IT solutions that support the business process is rapidly increasing. This implies that the enterprise architecture must be continuously maintained and governed. In addition, strategic developments in IT require architecture governance. Being an instrument to achieve strategic goals, the anchoring of the architecture work will normally move towards the top management.

32. Statistics Norway has also implemented best practice processes for project and portfolio management. Prince 2 is used as a framework for project management, which is centrally supported on the government level. The project proposals are systematically assessed according to user needs and sound business cases. We have also established a framework for Management of Portfolios, to ensure common prioritization and balancing of the project portfolio according to the available capacity. A project office has been created, and a role of Director of Project Portfolio has been defined at the top management level. However, we still need to implement sound roles and procedures on the level of clustering projects into programs. This should enable better coordination of projects. Currently the projects tend to become too big or too complex, occasionally leaving responsibilities on the project owners and projects managers that are beyond their capabilities.

33. One of the experiences has been that the project portfolio management processes can end up being too rigorous for some of the typical projects. The time spent to achieve a decision which must be taken outside the project can be unacceptably long, hampering the progress of work. The governance issued must be fit for purpose, and cover change management in general. This will be subject to

evaluation, where we will also look for ways to better categorise development work into more than one defined process as we have created for the big, strategic projects.

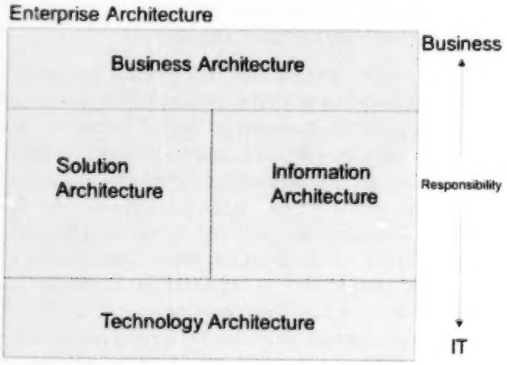
VI. Enterprise architecture

34. Statistics Norway has acquired valuable input to the modernization work from international collaboration. The development of international models like the Generic Statistical Business Process Model (GSBPM)[4] and the Generic Statistical Information Model (GSIM)[5] has given the necessary direction for national development work. This has led to increased efforts in building capacity in development and maintenance of a framework for enterprise architecture at Statistics Norway.

35. The first architecture discipline was established within the IT function. The need for common systems to support the life cycle of statistics production required the development of more complex systems, which made the systems more business critical. The systems and solution architects entered the scene. In addition, the requirements on IT to provide for best practice management of IT maintenance and operation were taken on board. However, we still lack sufficient capacity to meet the current needs.

36. The ability to provide for a coherent and holistic planning of the IT infrastructure and applications requires a corresponding action on the business side. The role and responsibility of an enterprise architect should be established, to take care of the complete architecture framework needed. The work on enterprise architecture is to some extent also encouraged by the international discussion on the industrialization of official statistics. The idea of harmonising the different parts of the business of official statistics according to common high level models to provide for collaborative development and sharing of tools and solutions has been brought further into the development of a common framework for enterprise architecture.

Figure 2
Simplified Enterprise Architecture Framework



37. International business models are used as reference models, and create the core of the business side of the enterprise architecture. GSBPM is the core business process model within the business architecture, just recently supplemented by the

[4] Generic Statistical Business Process Model: www1.unece.org/stat/platform/display/GSBPM
[5] Generic Statistical Information Model: www1.unece.org/stat/platform/display/gsim

generic activity model for statistical organizations (GAMSO)[6]. As mentioned before, Statistics Norway has developed a specific version of GSBPM, which is defined in three levels of detail, and is accessible on the intranet.

38. GSIM defines the high level of the Information Architecture, and the Common Statistical Production Architecture (CSPA)[7] outlines the principles and model framework within the Solution architecture. Recently, the need to agree upon a common Logical Information Model for statistical data as part of the Information Architecture has been launched. Such a model would become a reference for a logical model of statistical data in Statistics Norway. This approach to large and complex development projects has already been used in Statistics Norway, in a joint venture project to establish a scientific environment for on-line research on Norwegian register data, where the developments are guided by a comprehensive model for all the information objects handled by the system[8].

39. Well-defined Business Architecture and Information Architecture are tools to provide for a high-level specification for the design of the IT systems, whether they will be procured and implemented or built in-house. To some extent we therefore regard the Enterprise architecture as constituting a demand and supply side. In order to provide for better coordination and guidance of the strategic developments, Statistics Norway decided to move part of the architecture work out of the IT Department. As long as architecture work is inside IT, it is also regarded to be an IT "thing". There is a need to get the rest of the business and the top management on board the architecture driven development. While the Business Architecture is business owned and the Technology Architecture is owned and managed by IT, the business meets IT in common discussions, interests and competence within the Solution and the Information Architecture. The field of Information Architecture is a prominent part of the business competence in a statistical agency. CSPA is to some extent a technical response to the business process model, which also needs a consistent information model to be able to orchestrate IT components into the complete support of business processes.

40. The shared and coordinated responsibility between the business and IT concerning enterprise architecture is an initiative to provide for improved business driven developments. It allows the business to concentrate on defining what they need and where to go, leaving the responsibility on "how" to get there to IT.

VII. Initiatives to improve business agility

41. The Lean project has so far focused on continuous improvement within the current operational work. At the same time, the efforts to become more professional in the governance and management of big, strategic projects in statistics have led to a comprehensive framework for project portfolio management. However, the middle level of change management has not caught sufficient attention. This covers small and medium size development work, changes needed to encompass new regulations, medium complexity technical upgrades etc. The volume and impact of this work is perhaps highly underestimated. The way we are able to break down big development work into smaller pieces, could also bridge the gap between the way we run the

[6] Generic Activity Model for Statistical Organisations
www1.unece.org/stat/platform/display/GAMSO
[7] Common Statistical Production Architecture: www1.unece.org/stat/platform/display/CSPA
[8] RAIRD Information Model (RIM), Linnerud, J., Risnes, Ø. and Gregory, A. For information about the RAIRD project: www.raird.no

strategic developments and the continuous improvements and mid-size projects. The need for change has become normal; hence it should not always be treated as something special.

42. The need to improve the agile capability within the business is evident, but not specific to Statistics Norway or even to the statistics industry. However, it is neither possible nor desirable to change the business in all aspects and in all areas at the same time. While we professionally gather a lot of ideas and experience from the international work within official statistics, in general there are also lessons to be learned from other industries. Streamlining production processes and speeding up development processes is not specific to official statistics. Gartner, Inc. (www.gartner.com) is a leading information technology research and advisory company. Statistics Norway uses their analyses and advice for guidance and inspiration in some of the challenging areas within IT and business development.

43. Gartner describes the need to think about what they call Bimodal IT[9]. Bimodality represents a way to bridge the valuable elements and lessons learned from the industrialization of IT, with the capabilities needed to cope with the upcoming challenges of digitalization. This basically means how to combine risk management and internal focus on efficiency gains, with the need to become more innovative and externally focused, i.e. on new data and user needs.

44. Bimodal change processes within the business represent two complementary and valuable strands on how to develop the business, and how to run IT. On the one hand, we need to ensure that crucial parts and services within the office are solid, reliable and predictable. On the other hand, we need to change fast, to innovate and to experiment. In the figure below, some differences in characteristics are pin-pointed.

Figure 3.
Characteristics of Bimodal IT

Mode 1		Mode 2
Reliability	**Goal**	Agility
Price for performance	**Value**	Revenue, brand, customer experience
Waterfall, V-Model, high-ceremony IID	**Approach**	Agile, kanban, low ceremony IID
Plan-driven, approval-based	**Governance**	Empirical, continuous, process-based
Enterprise suppliers, long-term deals	**Sourcing**	Small, new vendors, short-term deals
Good at conventional process, projects	**Talent**	Good at new and uncertain projects
IT-centric, removed from customer	**Culture**	Business-centric, close to customer
Long (months)	**Cycle Times**	Short (days, weeks)

Source: Gartner

45. Agile development methods have been adopted in a number of development projects in Statistics Norway. The Lean projects contribute further to becoming more agile. Especially on the input and output side, we have recognized the need to speed up deliveries. This has led to a less rigorous need for managing change in

[9] Bimodal IT: How to Be Digitally Agile Without Making a Mess, Article from Gartner EXP Report 2014 No. 5 (www.gartner.com)

large, comprehensive projects. Instead, large and small improvements are continuously prioritized by the business and carried out by IT using for instance Kanban methods. In other areas, the consequence of change must be elaborated and controlled, providing for the use of more traditional plan-develop-run methodology.

46. The key objective is to establish change management throughout the organization, in a way that is fit for purpose in the types of change required. We try to develop a process oriented statistical production system to enable the statisticians to implement new statistical products in the production chain, without the need for IT expertise. When we succeed in this, new products and normal changes required from new regulation and new user needs, should be implemented much faster than today. However, we must also provide for speedy changes in the production system itself, and especially in areas where we expect changes to happen rather frequently.

47. One way to achieve this is to categorize the applications according to what Gartner defines as a pace layered model[10]. This is an approach that segments the application portfolio based on business value and rate of change. Gartner has defined three application categories or "layers" to distinguish these application types and to help organizations and IT to develop more appropriate strategies for their development and support.

48. These categories are described as:

(a) Systems of record — established packaged applications or legacy homegrown systems that support core transaction processing and manage the organization's critical master data. In statistical offices we should define standardized systems that support the business processes in this category, together with master data and metadata systems like statistical populations systems and master metadata systems. This category comprises applications that are subject to major change every 10 years or so;

(b) Systems of differentiation — consist of applications that enable unique company processes. Within the office we should regard this as being systems which are specific to certain products or areas of the business that cannot be accommodated only by systems of record. This comprises applications that typically are subject to change within a timeframe of 2-5 years;

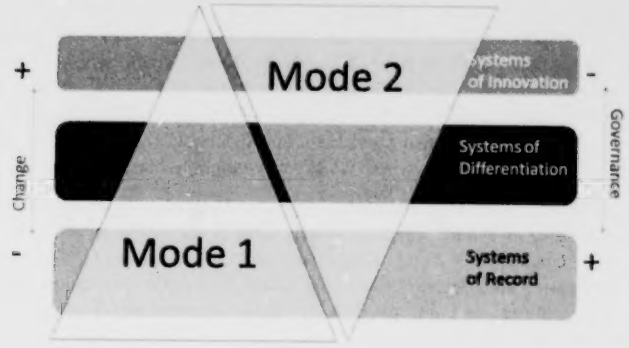
(c) Systems of innovation — new applications that are built on an ad hoc basis, often without full knowledge of needed requirements, and systems that are exposed to users or technologies where rapid change is required. In the office we typically could categorize (parts of) the systems on the input or output side in this area, where we typically address new business opportunities and needs. New demands and opportunities require change within 0-3 years.

49. Pace layering and bimodal IT are complementary. Bimodality orchestrates two distinct modes of capabilities, one for deploying services and one for solutions, which both relate and should be part of the corporate governance model for the entire IT function, including the business functions themselves. The balance between mode 1 and mode 2 will likely evolve over time. Mode 2 will start modestly, with a project focus; but as it expands, mode 1 will shrink. Once a mode 2 capability has been created, mode 1 involvement in systems of innovation will most likely diminish.

[10] Accelerating Innovation by Adopting a Pace-Layered Application Strategy, Jim Shepherd, Gartner Research (www.gartner.com)

50. Statistics Norway has just started to define processes which are more aligned with the agility defined by the nature of applications and areas of the production.

Figure 4.
Pace Layering and Bimodal IT are complementary.



Source: Gartner

51. Referring back to some viewpoints mentioned in the chapter on strategic governance, these strategic predictions from Gartner underpin the need to establish change as the normal mode within certain areas of the office. This means that we need to bring the Lean development and project portfolio development approach together. Again, this should be done under a coordinated, corporate governance and control.

52. If we look at international collaboration efforts like those encouraged by the High Level Group for the Modernisation of Statistical Production and Services (HLG), different preconditions and requirements could be assessed within the different categories. Just to mention a few:

- Cooperation in mode 1 (core systems) requires plan-driven and approval based governance, for instance also predictable maintenance and management of shared services and solutions. No such mechanisms exist today;
- Cooperation in mode 2 (systems of innovation) requires speed of development and time-to-market in deployment. Speed of (shared) development is a bottleneck within international collaboration today.

53. These challenges must be addressed, but they are not insuperable. It would be particularly worthwhile to join efforts in innovation. The effort to experiment with big data at the international level is a very good example of such an effort. Joint funding of common experiments will pay back, basically because no legacy systems or other typical obstacles that hamper cooperation exist.

VIII. Conclusions

54. Statistics Norway is striving to speed up the processes of change and development. Streamlining statistical production, quality reviews and Lean methodology alone are not enough – all such initiatives have to be followed up and guided in the same direction. To some extent the quality reviews themselves are a tool for such strategic guidance. The ability of the business to accommodate change is an important measure of quality.

55. The work processes for change and development of processes and products must be fit for purpose. The application portfolio should be assessed by its business value and the rate of change. Quality assessments and Lean methodology should ensure that we accommodate change management into the strategic governance at all levels. To speed up the pace of change is not only about sufficient funding and the right skills, but doing the right things the right way.

56. The top management at Statistics Norway has taken measures to ensure that the different initiatives are coordinated and move the institution in the same direction, in accordance with the strategy and capacities. Quality and improvement work is a continuous effort. User needs change over time, so does the environment for producing statistics including the technological possibilities. Consistency of purpose and management support on all levels is important. Change management will be further enhanced and implemented throughout the organisation, supported by strengthened capabilities to guide developments on all levels according to the overall business goals and strategies.

57. Concerning international collaboration, the vision of the High Level Group for the Modernization of Statistical Production and Services also requires the statistical offices to become more agile. International collaboration is useful in more or less all areas of our core business, which means that international collaboration also needs governance models that are fit for the purpose. We need to share efforts to create new value for the users, by participating in joint venture projects and partnering outside statistics. Harmonized business process models and enterprise architectures are a precondition for increased collaboration and shared developments. This calls for active engagement widely taken up by top management in statistical organisations.

United Nations ECE/CES/2015/38

Economic and Social Council

Distr.: General
8 May 2015

English only

Economic Commission for Europe

Conference of European Statisticians

Sixty-third plenary session
Geneva, 15-17 June 2015
Item 4 of the provisional agenda
Modernization of statistical production and services and managing for efficiency

Modernization of the statistical system in Poland

Note by the Central Statistical Office of Poland

Summary

The paper describes the efforts and plans of the Central Statistical Office of Poland in modernizing the statistical production process to increase its efficiency. To achieve this goal a number of activities accompanying organizational and coordination processes are or will be carried out as described in the paper.

The paper is presented for discussion to the first session of the Conference of European Statisticians' seminar "Modernization of statistical production and services and managing for efficiency".

GE.15-09215 (E)
1509215

Please recycle

I. Introduction

1. The current model of organization of surveys in the Central Statistical Office of Poland (CSO) continues the tradition and experience of the previous system of public statistics. This system was designed as a model known as "stovepipe", so that different statistical surveys are carried out for each domain and various organizational units are in charge of them. Approximately 20 departments and specific statistical offices take part in the survey process from the beginning to the end: from designing the survey to making the results available. The leader unit specialising in a given thematic field plans, carries out surveys and draws up results in accordance with prepared procedures, separate for each survey.

2. Eight main stages can be identified in the statistical production process: specification of needs, designing, building, data collection, data processing, data analysis, data dissemination and ex-post evaluation. The above-mentioned main stages of the survey process are basically in accordance with stages determined in the Generic Statistical Business Process Model (GSBPM) which presents the standard approach[1] *inter alia* to the description of the statistical production process. However, in reality the Polish statistical production process occurs separately for every survey and is executed in various ways, for example often with significant overlay of the boundaries between particular stages of the process.

3. With such an organization of statistical production, problems occur as a result of the silo effect. The separate organizational units are easily limited to their own tasks and attempt to implement them in their own way, but often trying to use common methods of work. As a result, the following problems may occur:

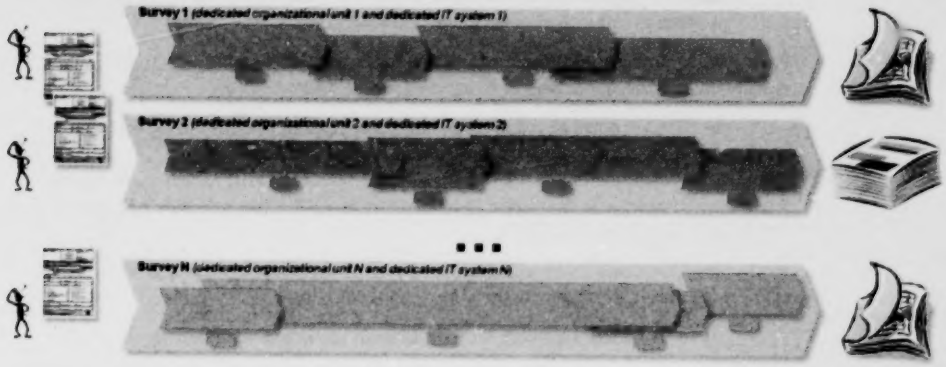
- High diversity (including lack of compatibility and standardisation) of work processes;
- Redundancy and often lack of specialisation in executed activities, in particular redundancy concerning the scope of information obtained from respondents;
- Duplication of competences being a result of, for example, lack of specialisation of activities;
- Development and maintenance of separate IT systems for each survey;
- Predominance of surveys for which data are collected via forms and questionnaires directly from the respondent with insignificant use of existing information available for instance in the public administration databases.

4. The above-described problems are visually shown in the figure 1.

5. Based on the above deliberations, it can be said that this type of organization of statistical production leads to having two hundred, mostly independent production lines, one dedicated for each statistical survey.

[1] www.unece.org/stats/gsbpm

Figure 1
Diagram representation of the current model of statistical production


Survey 1 (dedicated organizational unit 1 and dedicated IT system 1)
Survey 2 (dedicated organizational unit 2 and dedicated IT system 2)
Survey N (dedicated organizational unit N and dedicated IT system N)

II. Organizational change

6. The assessment of the organization of the CSO statistical production processes induced making a decision to introduce an integrated model of organizing the surveys by using the GSBPM rules. This is a desired direction of organizational changes of official statistics, not only in Poland. Problems arising within the current organization of statistical production in Poland are, to a significant extent, similar to those which appear in other organizations mentioned in the sector specific publications[2] as "*Autonomous Stovepipes*".

7. The introduction of a new model of organizing statistical work aims to optimise, re-design and standardise work processes. The change requires adequate preparation, and in the first place carrying out an analysis of previous work processes to identify similar activities with analogical results which require similar competences. Based on this analysis the GSBMP rules can be applied in a project to create more uniform and coherent standardised processes to replace the present production lines.

8. Changing the model of organization may influence many aspects of statistical production and, thus, requires many preparatory activities such as:

- Defining production processes in specific research realities;
- Integrating surveys in different domains of statistics – it is difficult to integrate all 250 topics of surveys carried out at present in the system of Polish statistics;
- Building a system of metadata as the basis for integration and standardisation;
- Preparing a concept and implementing a new integrated, information technology infrastructure for drawing up surveys;

[2] Bo Sungren, "Towards a system of official statistics based on a coherent combination of data sources, including surveys and administrative data", 2011

- Adapting to a change of the organizational structure that fits the integrated model of organization and supports process-oriented approach to statistical production (all as part of current practice) requires a lot of time.

9. When preparing for these organizational changes, a number of activities were already undertaken, for example:

(a) "A Concept of statistical surveys organization until 2020" was drawn up to be a model which provides direction to the activities of implementing the new organization of the statistical production process based on GSBPM and planning the schedule of implementation. The Polish model is called the Integrated Statistical Business Process Model (ISBPM);

(b) A Team to re-organise the system of statistical surveys and a few sub-teams to prepare specific processes were appointed;

(c) An ISBPS project was developed for implementation in the Polish statistics;

(d) Works on the System of Statistical Metadata as part of the Public Statistics Information System (PSIS) and adaptation of the information technology system to the new organization of the surveys are being carried out.

(e) The following priority tasks were selected for practical implementation:

- Examining user needs;
- Standardising and integrating data;
- Analysing the survey results;
- Improving communication with users and respondents.

10. To learn from the experience of other national statistical systems, CSO takes part in activities of the High Level Group for the Modernization of Statistical Production and Services and in committees appointed as part of HLG. CSO also analyses practical solutions used in various institutions and countries.

11. As part of the priority tasks realised, the following activities are carried out in the CSO:

(a) A methodology for diagnosing user needs, describing rules and ways of analysing user needs, and a communication strategy were developed. Specification of user needs is the first stage of the statistical production process. Adequate use of the analysis guarantees better adaptation of statistical products and services to user expectations. It will also make it possible to better adapt the scope of official statistics to the information needs related to existing and new social and economic phenomena. An assessment of the existing information resources will also make it possible to define gaps, identify any excess of information produced and avoid duplication of effort.

(b) A number of tasks were implemented to improve communication with stakeholders. Communication is an essential element of proper and effective functioning of official statistics. Due to the well-planned, organized and realised communication, CSO can effectively:

(i) Inform and encourage stakeholders and explain statistical products and services to them;

(ii) Promote products and services;

(iii) Maintain cooperation and good relations;

(iv) Educate users and disseminate knowledge of statistics;

(v) Eliminate any possible conflict situations (by predicting the effect of activities).

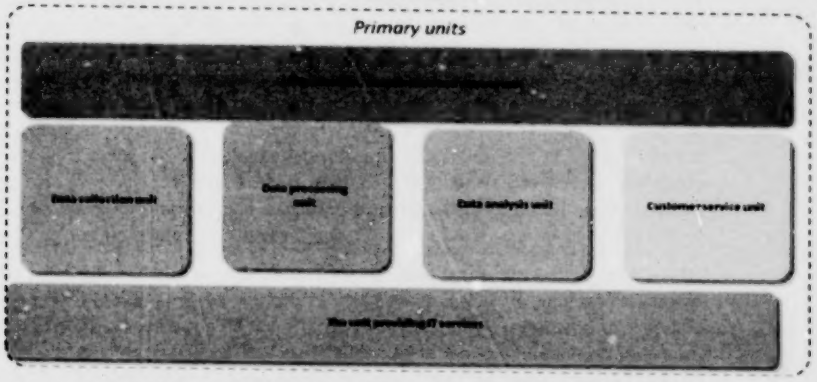
(c) A System of Statistical Metadata that brings together all metadata of variables will be created as a tool for inventorying and integrating the variables. An intention to modernize statistical production is observed in many countries, particularly concerning data collection. In the European Union, Eurostat stresses the need to increase effectiveness of data collection and reducing the burden on respondents. Existence of unjustified differences in the metadata which makes it difficult to integrate data collection both on national and European levels is universally noticed. Data and metadata should, thus, be comprehensively reviewed to standardise similar variables. Some of these should be possible to be integrated into one variable to be used for various surveys without great standardisation effort (e.g. by standardising formats, definitions, acquisition time limits, etc.).

III. Organizational structure and key tasks

12. Organizational structure is a relevant element of the ISBPM described in the previous chapters. It is also relevant to enable and support positive effects that it brings. According to the process description, CSO may identify basic units, such as:

- Control, coordination and organization Unit;
- Data collection Unit;
- Data processing Unit;
- Data analysis Unit;
- Customer service Unit;
- Information technology (IT) services Unit and other supporting units (dealing with staff management, administrative and finance management).

Figure 2.
Basic units of the Integrated Statistical Business Process Model



13. The basic assumption determining the multi-layer structure proposed herein is the concentration of similar tasks into task groups that require similar knowledge, experience and competence and form together a particular organizational unit.

14. One of the main tasks when planning the implementation of ISBPM will be to fully ensure the efficiency of carrying out statistical surveys. It is important to safeguard the integrity of each statistical process and ensure the appropriate competence for efficient coordination with all remaining basic organizational units. Therefore, the ownership of the process groups has been assigned to the *Control, coordination and organization Unit*. This unit is responsible both for the result of the production process as well as for defining and monitoring the measures of its efficiency, management of the process, defining the course of action etc.

15. A relevant support for the *Control, coordination and organization Unit* is provided by the CSO management which guards the process approach as outlined in the ISBPM. The management will provide the *Unit* with support in case of significant problems and issues requiring action on the level of the entire organization.

16. The fundamental task of the *Data collection Unit* is to carry out different data collections, on the basis of the necessary preparation by the *Data processing Unit*, using the most adequate methods of obtaining data. The products of this unit are the data collections gathered for the statistical products group and transferred to the *Data processing Unit*. This unit is the owner of the process of data collection.

17. The fundamental product of the work of the *Data processing Unit* is the micro and macro data repository appropriately prepared and fulfilled with data. One of the tasks is to analyse the needs of the *Data analysis Unit*. This need refers to the aggregates, the resulting micro data of statistical products and the processing of the obtained data into a data set to be used for compiling statistics and conducting analyses (i.e. value collections of variables collected for the statistical products groups). This unit is the owner of the process of data processing.

18. The main task of the *Data analysis Unit* is delivering to the *Customer service Unit* products based on the needs for statistical products. This task is implemented on the basis of a detailed data analysis, applying the statistical confidentiality principle and finally confirming that the data are made available. The fundamental result of the work of the *Unit* should be publication aggregates and the product groups statistically recorded in the analytical micro and macro data repository (data Warehouse). This unit is the owner of the data analysis process.

19. One of the tasks of the *Customer service Unit* is the preparation of statistical project products, on the basis of identified information needs (i.e. requirement specification regarding the products fulfilling the identified information needs of users). These statistical project products create the needs for statistical products. The second key task is making statistical information available in the form of final products prepared in the statistical production processes, including their transformation to the selected form for delivery, such as printed publication, electronic format etc. A relevant element of the *Unit*'s activity will also be the promotion of offered official statistical products. The *Customer service Unit* is the owner of the needs specification process and the data dissemination process.

20. The *Information technology (IT) services Unit* was classified as an organizational unit. That was mainly due to the role of IT solutions to support the activities of other units. Ultimately, that activity has an influence on the efficiency of the entire production process of official statistics. The *IT service Unit*, on the basis of requirements regarding the tools or IT services specified by all other units, prepares and provides IT tools and services. This *Unit* exercises, adjusts, tests and implements a set of tools and IT services, ensures continuity of work of the entire IT environment supporting the work of the other units. This *Unit* is the owner of the IT tool and services design process.

21. The proposed and above described structure exerts pressure to the organization of statistical surveys and the current functional specialization of the organization.

22. In order to proceed towards the new organizational status proposed herein for CSO, a gradual approach will be applied, i.e. step-by-step and stretched in time. Such an approach seems justified not only by the scope of changes but also by the fact that the proposed changes will require proper preparation and motivation of people engaged in this process. People constitute the fundamental capital of the organization, which, if appropriately organized and directed, determines the success of each organization functioning. People – with their knowledge and willingness to conduct creative operations – shape the organizational culture and formulate strategies and actions to implement them. The effective use of the existing competences and the development of new ones constitute a key success factor for the process of proposed organizational changes.

IV. Conclusion

23. There is no doubt that the generic model of organization of surveys is more effective from the point of view of:

- Managing the personnel resources, including adapting the competences to the needs of new organization of the surveys;
- Using efficiently the financial resources available for statistics;
- Improving the data quality as well as the functioning of the whole statistical institution;
- Organizing the process of statistical surveys and better using the potential of official statistics.

24. These are important reasons why decisions were made to change the organization of statistical surveys in Poland, while being aware that it is a many-sided undertaking which requires acceptance of statisticians as well as wide preparation by both the whole institution and its staff.

25. The implementation of the new model of organization of surveys constitutes a strategic decision bringing benefits for many years. Still, as in the case of every organizational transformation it creates certain costs. Therefore, the decision to implement ISBPM should naturally be preceded by a detailed analysis of the benefits and costs. However, with the benefits stemming from the use of the common solutions, a great potential seems to be underlying the proposed model. This model seems to eliminate numerous inefficiencies, including the following:

- A limited number of connected and integrated IT systems instead of hundreds existing;
- Several specialized organizational units, set up according to task specialization, concentrating specialist competences instead of many "independent" organizations;
- Standardized statistical survey processes, using common methods and best practices for various processes instead of using an independent and individual approach for each survey;
- Comprehensive review of all conducted surveys: eliminating sources of structural ineffectiveness by grouping products;

- Instead of expensive data, collected from direct surveys, obtain cheaper data by using alternative sources like administrative sources or Big Data, and by sharing data within the organization from surveys in different fields.

United Nations ECE/CES/2015/39

Economic and Social Council

Distr.: General
4 June 2015

English only

Economic Commission for Europe

Conference of European Statisticians

Sixty-third plenary session
Geneva, 15-17 June 2015
Item 4 of the provisional agenda
Modernisation of statistical production and services and managing for efficiency

Building a common production environment for the statistical production process

Note by Statistics Sweden

Summary

This paper gives a short overview of the progress of Statistics Sweden in building a common statistical production environment. Furthermore, the paper describes the ongoing work and elaborates on some of the lessons learned during the process of introducing a common statistical production process.

The paper is presented for discussion to the second session of the Conference of European Statisticians' seminar "Modernisation of statistical production and services and managing for efficiency".

GE.15-08985 (E)

1508985

Please recycle

I. Introduction

1. Since 2008, Statistics Sweden has moved to a process-oriented approach in statistical production. The aim is to make production more efficient and standardised, using common methods and tools wherever possible. One important aspect of this is building a common information technology (IT) platform where existing and new tools can communicate efficiently with each other, and where statisticians can choose the necessary tools for each survey. For data collection, this platform has been realised through the so-called TRITON project, which is now used by many surveys to collect and validate data through web collection and scanned paper forms. The structure set by the TRITON project is now being applied and expanded to other parts of the statistical production process (processing, analysis and dissemination). The paper describes how far Statistics Sweden has come, describes the ongoing work and elaborates on some of the lessons learned so far.

II. The current situation and development so far

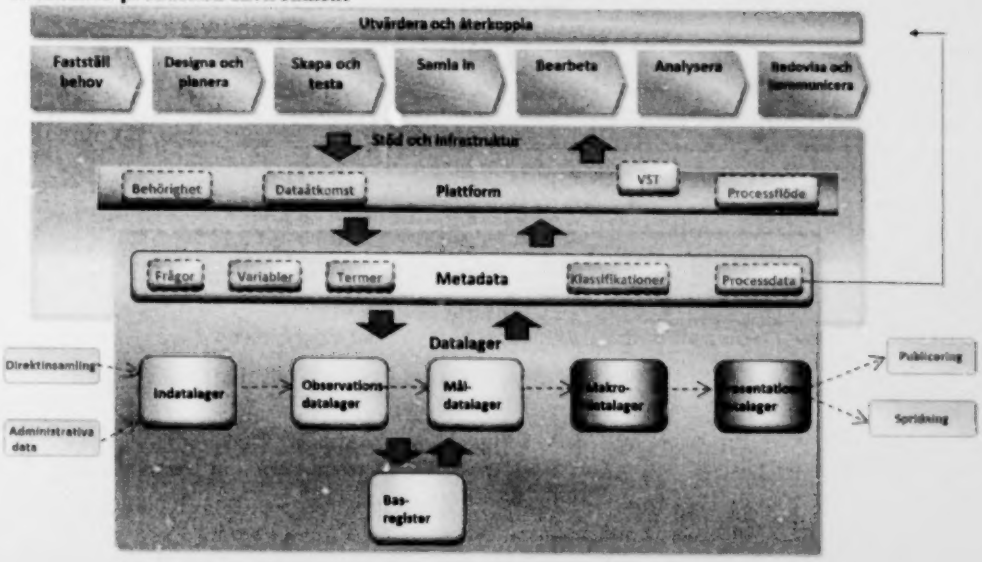
2. In 2008, Statistics Sweden introduced a process-oriented approach to the statistical production process, introducing our own version of the Generic Statistical Business Process Model (GSBPM, see Figure 1 below). The adapted GSBPM was divided into five sub-process groups, which were all given a devoted process owner (and a deputy). The role of the process owner is to lead the development and maintenance of methods, tools and routines within his or her sub-process, to give support and to provide guidelines and descriptions of the process and how to run it.

3. The descriptions and guidelines on how to run the statistical production process were put on the internal website of Statistics Sweden in a tool called the process support system (PSS). PSS can be seen as a handbook in statistical production at Statistics Sweden, and includes all agreed-upon common methods, IT tools, routines, checklists, templates etc. that the different surveys must use in their individual production processes. But the vision for PSS has from the start been to create something that is even more active, giving each survey more specific information based on the characteristics of the particular survey.

4. At the same time, there was a need for modernisation of many of the survey specific IT tools used – these being a result of the decentralised organisation we had and the need for development around the millennium shift where a personal computer (PC)-based platform was introduced and many systems needed to be replaced quickly. This resulted in many local systems. Now, around 10 years later, the drawbacks of these systems were becoming clear – dependency on specific persons, old technical solutions that were no longer in use or supported, and expensive maintenance. The drawbacks were especially notable since there were new, more common solutions being built (for example for web data collection) that had to be integrated with all the survey specific systems. So the idea was raised to combine the initiatives and start building a common production environment with the parts outlined in the conceptual picture in Figure 2 (in Swedish). With the production process at the top and a new, structured data storage function at the bottom, a number of IT tools or services were envisioned to access the data and enhance it, with metadata being used to connect the tools and the data. The IT tools are accessed through a new version of the PSS (or VST in Swedish, as can be seen in the picture).

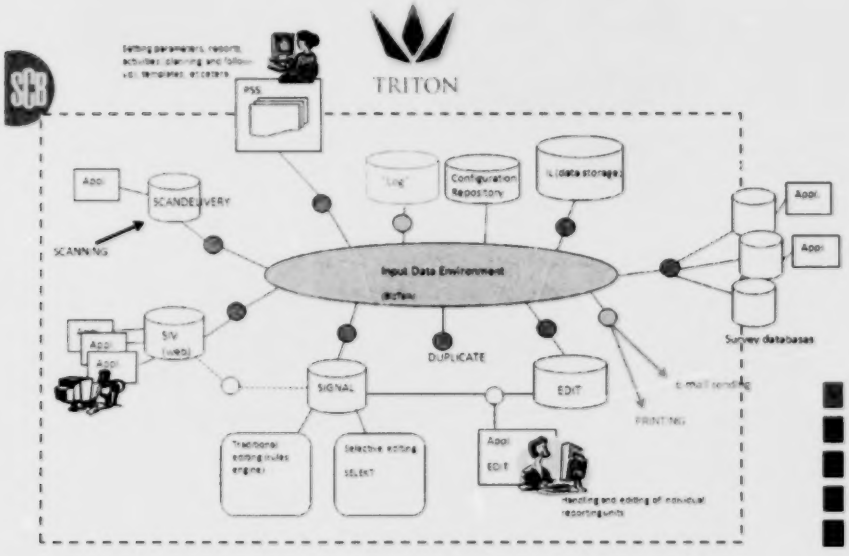
Figure 1
The Swedish adaptation of GSBPM

Figure 2
A common production environment



5. It was decided to start the development at the data collection part of the production process. There were several reasons for this, the main one being that many systems were outdated and that with data collection being a staff-intensive part of the production process, the potential gains in efficiency were large. Also, the integration with our standard tools for web collection, scanning etc. was not very efficient, and there was a need to eliminate manual handling in the production to minimise the risk of errors. Furthermore, we had a new tool for selective editing (SELEKT) which needed to be integrated with an environment for investigation of potential errors. A project was initiated to build a common production environment for data collection that could fulfil these aims, and that could also lay the foundation for a common production environment on a larger scale. This project was TRITON. The first version of TRITON was launched in the middle of 2011, and since then a number of projects and releases with further development have been carried out, also in many ways taking the gradual implementation of surveys in TRITON into account, and re-building parts and infrastructure where necessary. At the moment, Statistics Sweden is already using version 4.8 of TRITON and has around 30 surveys in the system, with another 20 due to be developed before the first quarter of 2016.

Figure 3
TRITON



6. The main parts of TRITON can be shown and listed as follows:

- The Input Data environment that steers the information flows between the different parts
- The data storage (IL) which has been re-created as "shards" or identical databases for each survey and production round to gain independence and performance
- Connections to data coming in through web data collection and scanning – data collection through telephone interviews
- Editing through an error localisation service (SIGNAL) and a connection to SELEKT

- An application (EDIT) for treatment of reporting units – collection, investigation of potential errors, communication etc. This tool has been enhanced greatly since the first version, including several ways to look at data, handling changes in reporting units and more
- Printing and sending of letters and e-mails to reporting units
- Delivery of data to other systems
- And – last but not least – the new version of PSS called "Process areas".

7. Process areas are a kind of dashboard to look at production. They give access to all the tools needed, contain process data reports on the current status and include a quality assurance and control system. The process areas and the overall architecture of TRITON are the tools that can be re-used in other parts of the common production environment. The process areas are built in a hierarchical structure with three levels: Survey, Production round and (for data collection) Collection round. In parallel to Collection rounds we envision two more process areas at the lowest level: Calculation/Processing round and Dissemination round (these names are preliminary and subject to change). Introducing Process areas is also taking a large step towards an active PSS, since the content of a Process area for a specific survey is dependent on the design choices of that survey. This is achieved by letting the survey fill in a questionnaire detailing its main design choices, after which the area is generated with content based on the information in that questionnaire.

8. In 2012, discussions started on where to go next in expanding the common production environment. There were two main options – either to continue with the processes coming after collection and micro editing (further processing and analysis) or to take the next step in the dissemination area where the process is already very functionalised (as is data collection) and where it would probably be easier to move forward. Since the IT tools for dissemination were also undergoing changes, some being modernised and others being replaced, it was decided to start in that area and to begin with a pre-study in the Process and Analysis area, where the way forward was a little more uncertain, and where a recently-started project dedicated to a number of surveys in short-term economic statistics could be used as a first step to see how to proceed.

9. The project in the dissemination area started in 2013 with identification of the main IT tools that would be replaced, and with discussions about how the process area structure would function for dissemination. It was agreed that the structure used for data collection would fit for dissemination as well. However, it was decided not to proceed immediately with introducing process areas for dissemination, but rather to focus on replacing the most important IT tools first. In 2014, the metadata structure for dissemination was put in place in the new tool "PX Meta", and in 2015 work is ongoing with a tool for planning publication on the Statistics Sweden website. Later this year some work on process areas might be undertaken.

10. In the Process and Analyse area, the project for short-term economic statistics has developed a tool for planning and running calculation, imputation, estimation and more. The tool is called KLON, and while not built to be a completely generic tool, it has the potential to become one. A bridge has also been built between TRITON and KLON so that data collected in TRITON is transferred automatically to KLON. But metadata-wise, KLON and TRITON are not fully integrated, meaning some set up of metadata is duplicated. All micro data editing is done in TRITON (with the EDIT tool). The first version of KLON was finished in early 2015. A few more products have already decided to move their calculations into KLON. This has meant that the planned pre-study on how to proceed with the common production environment for Process and Analyse was carried out on time and started late 2014 with the aim of describing how much of KLON could be re-used when expanding to a more generalised production environment. While conclusions

were a bit difficult to draw given that the pre-study was done before KLON version 1 was finished, it can be seen that several things from KLON will be very re-usable. It also became clear that there were somewhat differing opinions on how far it was possible to go with a common production environment for Process and Analyse, and which were the most important things to develop, so further internal discussions need to be held on this matter.

11. At the end of 2014, with all the different initiatives ongoing and with old systems still needing to be replaced, a new approach to the project was taken. An initiative to merge all the different parts of development into one project, also incorporating implementation in a number of surveys and replacing their old systems, has been taken in 2015. This large project is still being set up and organised at the time of writing, and will be fully operational in the second half of 2015, with the objective of leading the way to a common production environment for all steps of the production process from data collection to dissemination in 2015, 2016 and beyond. Beside IT development and implementation, a third part of this project will deal with the overarching structures, including architectural issues, common metadata (that are to be used across processes) and structures for the PSS process areas, among other things. Examples of common metadata are data on surveys, production rounds etc., variables and value sets including classifications, data deliveries and which survey uses which data. This is done to ensure that, in spite of the likelihood that we will opt for three different parts (data collection, processing and analysis, dissemination) within the common production environment, which may differ in some aspects, the three parts will fit together to eliminate duplication of metadata etc. To ensure this, the Generic Statistical Information Model (GSIM) will be used as much as possible. A comparison between GSIM and the information model used in TRITON has been carried out, and has not found large differences, but has rather pinpointed the need for a common understanding of the GSIM objects in the different processes in order to have an integrated environment while allowing each part to be built to satisfy the needs of the specific process. Much important common metadata is found in the "blue" business part of GSIM.

12. The TRITON architecture has also been used in a proof-of-concept project within the UNECE initiative of the Common Statistical Production Architecture (CSPA), successfully receiving answers to interviews from the computer-assisted survey processing, BLAISE, and sending them through the platform. This has strengthened the feeling that the architecture upon which our environment is built is strong and expandable.

III. Reflections: Issues and challenges

13. Building an IT platform for all parts of the statistical production process takes a lot of effort, resources and time. In the meantime, there are many other challenges that need to be met within a statistical office. The path Statistics Sweden has taken is based on taking one process at a time, and also working in three different areas in rather different ways. While the process is moving in the right direction at a steady pace, there have also been challenges.

14. The three main areas in the production process are now differently organised. At Statistics Sweden, staff in charge of data collection and dissemination are to a large extent centralised, whereas processing and analysing is mostly carried out by survey managers and methodologists survey by survey. Collection is highly dependent on IT tools to get an overview of the status of production. Dissemination on the website also uses many IT tools that are common across all surveys, while in process and analyse there are several tools that are useful but which are mostly used in combination with survey specific adaptations. This in turn is partly due to surveys having very different data storage solutions, i.e. there is not a common data warehouse for all surveys. This has meant that there is a need to take these considerations into account when building the common production environment; all parts

should not work in exactly the same way. On the other hand, a survey manager wants to have an overview of the production from start to finish, so the environment needs to incorporate that need as well. The main challenge is to make each process feel part of an integrated whole while supporting the specific needs of different processes.

15. Statistics Sweden has worked by development in stages, both between processes (data collection, then dissemination, then processing and analysing) and within processes (TRITON has been developed in stages and is still not "finished"). This has both pros and cons. The biggest pros are that results are revealed step-by-step rather than having to wait for a "big bang" when everything is finished, and that experiences from early implementations can be taken into account in further development. In fact, we strongly believe that waiting for a complete new system for all areas of production to be finished before releasing it would probably mean it never being finished at all. The biggest con is probably that one has to work with temporary solutions for a long while, combining the new common environment with parts and systems from older environments. Furthermore, the first versions released may be deemed "not good enough" even though further development is already planned and ongoing. Also, there is often a wish from different parts of the organisation to proceed fast in their specific area, making prioritisation difficult.

16. Moving from survey specific systems to a common production environment is a large step, and a common environment necessarily becomes a compromise between different demands from surveys used in very different production systems. There is a need to find solutions that are good enough for most while not moving all possibilities from all old systems into one big system instead, in order to make the new solution easy to maintain and to permit high performance. On the other hand, the new environment needs to be flexible to accommodate different kinds of surveys, not making production unnecessarily difficult for some kinds of surveys. Finding how much functionality to incorporate in an optimal way is a great challenge. The development and the strategy also needs strong support from upper management all across the office, because if asked to choose between having a completely custom-built survey specific system and joining a common environment, many survey managers would choose the first alternative if they had the economic possibility to do so. If the office wants as many surveys as possible to use the common production environment, it must be supported even if every demand from every survey cannot be perfectly met.

17. Implementation of the new environment is a fundamental step that might be easy to overlook. Surveys need to start implementing early to get a feel for the new system and to be ready in time for production. On the other hand, changes and adaptations take place right up to the last minute. Support in the implementation process is of utmost importance, and very different to IT development, so it needs to be led by other people than the developers. Also, post-implementation support is very important when production is running. In both these areas we have struggled a little to meet the needs of the surveys. There is also an in-built conflict between further development (adding new functionality) and maintaining what is already there, since the same staff and resources are needed for both. So we would advise any country that is going on the same journey to consider this very carefully and to devote necessary resources for it. In our case, it is also a journey from the small survey specific systems to large cross-office systems, with the inherent risk of being a "single point of failure". Resources must be devoted to the large systems because so many people are dependent on them working all the time, which is a different situation to when each survey has its own system.

18. Even though we have faced a few challenges, encountered some obstacles and progress has been a bit slower than anticipated when we started, we are still convinced that this is the right way to go. The updated strategy for Statistics Sweden ("Strategy 2020") also clearly points out the common production environment as an important means to

achieve our goals of producing high quality statistics based on the needs of our users, while easing the burden for respondents.

United Nations ECE/CES/2015/40

Economic and Social Council

Distr.: General
1 June 2015

English only

Economic Commission for Europe

Conference of European Statisticians

Sixty-third plenary session
Geneva, 15-17 June 2015
Item 7 (a) of the provisional agenda
Reports on the work of the Conference of European Statisticians, its Bureau and Teams of Specialists

Main directions of work in preparing a Guide on Measuring Human Capital

Note by the Secretariat

Summary

This document presents the outline and the main directions of work in preparing a *Guide on Measuring Human Capital*. The work is being carried out by a Task Force on Measuring Human Capital which was established by the Conference of European Statisticians in 2013.

The Conference is invited to take note of the main directions of work and the outline of the Guide, and send written comments and suggestions to the UNECE secretariat (andres.vikat@unece.org).

GE.15-11046 (E)
1511046

Please recycle

I. Background

1. Understanding human capital is of significant interest to policymakers. Statistics on human capital may help to understand the drivers of economic growth and the functioning of the labour market, as well as to assess the long-term sustainability of a country's development path.

2. Based on the outcome of a Conference of European Statisticians (CES) seminar in 2011 and the recommendations in the subsequent stock-taking report, CES established the Task Force on Measuring Human Capital in 2013. Its objective is to pursue the conceptual development of human capital measurement, with priority on developing experimental human capital satellite accounts.

3. The Task Force has 19 members: from Australia, Canada, Italy, Netherlands, New Zealand, Norway, Poland, Slovenia, United Kingdom, United States, OECD, UNECE, the Wittgenstein Centre for Demography and Global Human Capital, Universities of Maryland and Madison-Wisconsin, and the Central University for Finance and Economics in Beijing. Ms Ann Lisbet Brathaug from Statistics Norway chairs the Task Force.

4. The Task Force commenced its activities in 2014 towards developing a *Guide on measuring human capital*. At its first face-to-face meeting in June 2014 in Geneva, it agreed on the outline of the Guide and distributed responsibilities among its members for drafting the chapters. The drafts were discussed electronically and at audio conferences. At its second face-to-face meeting in April 2015 in Paris, the Task Force decided on a number of specific methodological issues and a timetable for the next steps. UNECE plans to engage an experienced editor to streamline and edit the various chapters of the Guide.

5. The Task Force plans to present a draft Guide to the February 2016 meeting of the CES Bureau. Upon the Bureau's approval, the Guide will be sent for consultation among all CES members in spring 2016, and subsequently submitted to the 2016 CES plenary session for endorsement.

6. The Conference is invited to provide comments and suggestions on the main directions of work presented below and the outline presented in Section III (please send the comments to: andres.vikat@unece.org).

II. Main directions of work

7. The *Guide on measuring human capital* will contain recommendations on the methods and sources to compile human capital measures. It will also discuss the setup of a satellite account for human capital, but concludes that more research is needed before this approach could be recommended. Therefore, as a first step, the guide suggests starting with the development of a satellite account on education and training.

8. The Guide will discuss the cost approach and the lifetime income approach as the main methods to measure human capital. Both methods combine the different aspects of human capital in a single metric, but give different values. The indicator approach will only be recommended for providing complementary information for volume estimates, quality adjustment and cross-checking. The Guide will not recommend the use of indirect methods.

9. The cost approach is in line with the Perpetual Inventory Method that is often applied for the measurement of fixed assets in the System of National Accounts

(SNA). It also provides an estimate of the resources invested in education and other human capital related expenses, which can be useful for cost-benefit analyses. This approach is also in line with the recommendations for the compilation of a satellite account for education and training. For the cost-based approach, the Kendrick methodology[1] will be introduced.

10. In the case of the cost approach, the Guide will mainly consider the following elements:

(a) Formal education services paid by households or provided for free or at reduced prices by government;

(b) Training and courses provided by employers and;

(c) The foregone earnings of students (time spent studying).

11. The last component will extend the SNA production boundary. Further work is needed to define its exact coverage. Taking into account issues related to data availability and measurement, other non-market activities will be kept outside the scope of these human capital estimates.

12. The lifetime income approach is more consistent with economic theory and better represents the relationship with productive capacity needed for future production. It also allows outputs to be measured independently of inputs when estimating the productivity of the education sector. The presentation in the Guide will rely on the Jorgenson and Fraumeni method[2]. For international comparability, the Guide will recommend using:

(a) The SNA concept of compensation of employees, supplemented with an imputation for labour income of self-employed and;

(b) Working age population without suggesting a specific upper age limit, as this depends on the actual retirement age in the country.

13. The Guide will advise countries to provide sufficient metadata to allow cross-country comparisons.

14. From a conceptual point of view, the lifetime income approach is more consistent with market price valuation of the assets. However, from the point of view of data availability, the Guide will encourage the development of cost-based measures as a country's first step in human capital measurement.

15. The Guide will discuss compiling both monetary and volume measures of human capital. The methods to arrive at volume estimates will be made conceptually consistent with general methodology for measuring volumes in the system of national accounts.

16. The Guide will outline a two-step approach to linking human capital estimates to SNA. At a first stage, it will recommend developing a satellite account

[1] Kendrick, J.W. (1976), The Formation and Stocks of Total Capital. New York: Columbia University Press.

[2] Jorgenson, D. W. and Fraumeni, B. M. (1989). The accumulation of human and non-human capital, 1948-1984. In R. E. Lipsey and H. S. Tice (Eds.), The Measurement of Savings, Investment and Wealth (pp. 227-282). Chicago, IL: The University of Chicago Press. Jorgenson, D.W and B.M. Fraumeni (1992), The Output of Education Sector, in Z. Griliches (ed.) Output Measurement in the Service Sectors, The University of Chicago Press, Chicago. Jorgenson, D.W and B.M. Fraumeni (1992), "Investment in Education and U.S. Economic Growth", Scandinavian Journal of Economics, 94, Supplement, 51-70.

for education and training. Such an account will use the supply and use framework as a starting point, introducing further breakdowns of the various expenses related to the formation of human capital by type of product and producing unit, including breakdowns into the various types of (formal) education. It will also recommend a breakdown of labour input and compensation of employees by level of educational attainment. This satellite account will provide most of the components of the cost-based method. The idea behind the satellite account is to describe the educational system from an expenditure perspective, i.e. which education services are provided and by whom, and who is benefiting from the services.

17. As a further step, the Guide will present a complete satellite account for human capital where expenditures on human capital are classified as investments. It will outline the different imputations in the accounts and the changes to the components of the production accounts (including Gross Domestic Product (GDP)), and the current and capital accounts. The reclassification of education and training expenditures as investments will obviously have an impact on GDP. The impact, however, is depending on the method and the assumptions used for estimating the investments.

III. Outline of the Guide

A. Chapter 1: Introduction

18. Chapter 1 will provide the context to the production of the Guide and explain its importance and policy relevance. It will then provide a short overview of each chapter.

B. Chapter 2: Concepts and definitions

19. Chapter 2 will provide a general framework of the concepts underlying the measurement of human capital and on the ways how human capital can be incorporated into the SNA. It will describe the two preferred methods to estimate the economic value of human capital: the cost approach and the lifetime income approach. The chapter will debate the appropriateness of fully integrating human capital in the framework of national accounts, particularly with the challenges posed by the measurement of the accumulation and obsolescence of human capital, and suggest the production of satellite accounts for education and training as the first step. The chapter will also elaborate on the differences between the cost and lifetime income approaches.

C. Chapter 3: Methodological issues

20. Chapter 3 will look at the methodological challenges related to the measurement of human capital. It will look at the theoretical basis for the cost-based and lifetime income approaches and their challenges, and highlight the need for arriving at estimates of stocks, investments and consumption of human capital for integrating human capital into national accounts.

21. The chapter will recommend inclusion of the time used for studying (foregone earnings) as a component of the cost approach, making clear that this estimate extends the production boundary of the current system of national accounts.

The chapter will recommend compiling both current price and constant price (volume) measures of human capital.

D. Chapter 4: Implementation and measurement issues

22. Chapter 4 will follow the methodological discussions from chapter 3, focusing on operational aspects of arriving at internationally comparable measurement and data. The main text will discuss the cost-based method, the lifetime income-based method and the indicator approach.

23. The data requirements for the implementation of human capital measures will also be discussed, including information on the methodologies and data sources for estimates on:

(a) Expenditures on formal education and on-the-job-training;

(b) Foregone costs for students in education and employees in training;

(c) Depreciation rates related to the various investments in human capital;

(d) Initial human capital stocks and;

(e) Price indices.

24. While data for (a) can be relatively easily obtained, the Guide will consider the necessary assumptions to calculate (b)-(e).

25. For the lifetime income-based method, the Jorgenson-Fraumeni methodology will be introduced. This chapter will also describe the necessary data and assumptions to estimate lifetime income, which includes:

(a) Educational attainment;

(b) Enrolment;

(c) School duration;

(d) Employment;

(e) Labour compensation;

(f) Income growth rate and discount rate;

(g) Volume index;

(h) Migration; and

(i) Training, including the relevant breakdowns by sex, age and education level.

E. Chapter 5: Satellite account for education and training

26. Chapter 5 will present a satellite account for education and training, which is the recommended starting point for the estimates of human capital. The chapter will show an extended supply and use table integrating various more detailed data on education services, time used for studying, and labour inputs. The chapter will present the recommended level of detail and identify the data limitations.

F. Chapter 6: Going beyond SNA

27. The chapter will provide a satellite framework for the full and consistent introduction of human capital into the SNA. The main concept is that economic aggregates such as GDP, investment, consumption, savings and wealth, change when the production and asset boundary in the SNA is broadened to include human capital.

28. The chapter will discuss the integration of both cost-based and income-based estimates of human capital into SNA, their impacts, technical components and necessary assumptions. Moreover, it will discuss the changes in the various accounts of the SNA, including its impact on GDP, gross national income (GNI) and household accounts more generally. The chapter will also indicate that this type of analysis can be expanded to include other types of human capital investment, such as investment in health, informal training and costs of child rearing.

G. Chapter 7: Human capital country studies

29. This chapter will showcase some practical examples of how human capital has been measured in some country-specific contexts, the type of data used and the possibilities of replication of these studies to other countries.

United Nations ECE/CES/2015/41

Economic and Social Council

Distr.: General
22 May 2015

English only

Economic Commission for Europe

Conference of European Statisticians

Sixty-third plenary session
Geneva, 15-17 June 2015
Item 4 of the provisional agenda
Modernization of statistical production and services and managing for efficiency

Product and service innovation at Statistics Netherlands

Note by Statistics Netherlands

Summary

This paper describes the strategic initiatives taken by Statistics Netherlands to increase the public value of official statistics. In the climate of budget cuts in almost all public organizations, it is of particular importance and requires continuous improvement and innovation of products and services. The paper describes how Statistics Netherlands strives to deliver factual information at the right time for evidence-based policy making. This strategy has implications both for the statistical work at Statistics Netherlands and for the way the communication is organized.

This paper is presented for discussion to the second session of the Conference of European Statisticians' seminar "Modernization of statistical production and services and managing for efficiency".

GE.15-10383 (E)

1510383

Please recycle

I. Introduction

1. Statistics Netherlands faces severe budget cuts, just as almost all public organizations in the Netherlands. In this climate, it is of particular importance to maximize the public value of official statistics. This requires continuous improvement and innovation of the products and services. An important goal of a national statistical institute is to enable rational evidence-based policy making, i.e. policy making based on facts. However, it becomes more and more difficult to characterize complex phenomena in society using just one or a few indicators. To overcome this problem, Statistics Netherlands is revising its strategy by reporting more underlying indicators in a coherent way to create a deeper and more complete understanding of developments in society. In this way, Statistics Netherlands strives to deliver factual information at the right time which can be used effectively by policymakers. This strategy has implications both for the statistical work at Statistics Netherlands and for the way the communication is organized.

2. The most recent Strategic Statistical Programme[1] of Statistics Netherlands covers years from 2014 to 2018. It has been prepared on the basis of extensive discussions with the main users. On the basis of these discussions, the Programme identifies the following as statistical priorities:

(a) New statistical products about *globalization* and internationalization of enterprises;

(b) New statistical products about the *labor market*, in particular about flexibilization and migration;

(c) New statistical products about *financial assets and liabilities* of households and enterprises;

(d) Better *accessibility* of statistics, in particular on the internet;

(e) More *specialized statistical studies* for central and local government;

(f) More *timely* statistics.

3. Due to budget cuts, resources for these priorities are very limited, at most two percent of the annual budget of Statistics Netherlands, and must partly be realized by dropping some statistical programmes.

4. In 2015, some further elaborations have been made. The first one is that Statistics Netherlands will focus much more on communication, presenting itself to users as a news organization.

5. Secondly, focus will shift from individual statistics to coherent statistical information describing the whole of a phenomenon in society; e.g. communication and statistics around unemployment focus now more on the phenomenon of the labor market as a whole and less on unemployment alone. This approach will also

[1] Every four or five years Statistics Netherlands draws up a Strategic Statistical Programme (Statistics Netherlands, 2013), which gives guidance to new statistical products and new ways of communication, and which is accompanied by a multi-annual budget proposal. This Program and the accompanying budget have to be confirmed by the Council of Ministers and are sent for information to Parliament. The separate years covered by the Programme are translated by Statistics Netherlands into annual work programs and by the responsible Minister into annual budgets.

identify gaps in the statistical program around this phenomenon as well as possible linkages between phenomena.

6. And thirdly, we have established an innovation unit in order to speed up the implementation of new ideas and tools, so that we may respond more quickly and flexibly to new user needs.

II. Strategic priorities for 2014-2018

A. User needs

7. Statistical priorities should be based on user needs. However, because statistics are a public good, it is not always easy to elicit user needs and user preferences. Therefore, Statistics Netherlands has an extensive network of advisory committees and contacts with users. The most recent Strategic Statistical Programme has been prepared on the basis of extensive discussions with our users or their representatives: ministries, organizations of provinces and municipalities, employer organizations, labor unions, press and universities, as well as with focus groups of selected users, such as young people and scientists. We feel that it is important to make this consultation as broad as possible, and to include diverse groups of users and also non-users.

B. Dropping statistical programs

8. Whereas users are always willing to propose new statistical programs or new ways of communication, they often have difficulties in setting priorities, in particular with reference to existing programs. A consultation round similar to that for new priorities has been attempted, but it delivered few concrete results.

9. However, because of severe budget cuts, it was necessary to drop some statistical programme in order to make room for new priorities. The alternative of no new statistical programmes was deemed unacceptable: the national statistical office would become increasingly less relevant if it was not responsive to changes in society and changes in user needs. So instead, we have followed a different approach where the balancing of statistical programs was made by Statistics Netherlands on the basis of criteria, referring amongst others to usage, costs, and external requirements; it is described more fully in Zeelenberg et al (2010). Of course, this list of programs to be ended has been discussed with the advisory committees and with ministries, but this has not led to major changes in the list.

C. New statistical topics[2]

10. The user needs as expressed in the consultation round were on the surface very diverse and very specific. However, they could be grouped into six categories:

[2] The text of this paragraph is mostly quoted from the Strategic Statistical Program.

1. New statistical products about globalization and internationalization of enterprises

11. As a result of globalization and digitalization, geographical features (e.g. national borders) and physical distances have become less important. Trade via the internet means that national borders now hardly play a role for consumers. While the relevance of international economic developments is evident for an open economy like the Dutch one, Dutch competitiveness and the appeal of the Netherlands as a business base is increasingly a regional rather than a national issue. This means that national figures are more and more often an average of very different underlying regions, each with its own strengths and weaknesses. The value of national statistics is thus reduced, increasing the need for figures at a regional level. In the next few years, new regional figures will be developed along these lines.

2. New statistical products about the labor market, in particular about flexibilization and migration

12. The labor market is becoming increasingly fragmented, as a result of which the distinction between employee and self-employed is becoming less clear, and new groups have come into being such as self-employed people without employees and pay-rollers. New 'groups' on the labor market and the underlying motives of the labor market actors concerned will be described. Insight into these aspects will provide policymakers, researchers and society with a more complete and better insight into how the labor market works.

13. Also, because of the lively public debate about international migration, it is important to chart the structure, flows and characteristics of labor migration and its consequences for the Dutch labor market and the social security system.

3. New statistical products about *financial assets and liabilities* of households and enterprises

14. New statistics will give as complete a picture as possible of the financial position of various groups in relation to their status on the housing market, life cycle, economic position and labor market status. This will make it clear, for example, how the rapidly expanding group of self-employed that does not have any employees covers their financial risks and the extent of their financial safety nets.

4. Better accessibility of statistics, in particular on the internet

15. The accessibility of data in our online database StatLine will be increased further in the next few years. The search engine and navigation structure of the website will also be improved.

16. In addition there is an increasing demand for open data. All our online data is now accessible as open data, i.e. through well-defined meta data and a common API; see http://www.cbs.nl/nl-NL/menu/cijfers/statline/open-data/default.htm and http://opendata.cbs.nl/dataportaal/portal.html#_la=en. Public and private organizations will then be able to develop themselves products, services and applications on the basis of our data.

17. Local authorities in particular have a great need for data at the local levels of districts and neighborhoods. We will make existing information – that can often be presented geographically – more accessible, and will expand the available information. Another development in this respect is the European harmonization of geographical data, making it easier to compile statistics on border-crossing regions.

18. Finally, new cryptographic methods such as secure multi-party computation (Wikipedia, 2015), might make it possible to combine advanced security and increased access to microdata. But this is an area that requires much further research, such as in the development of appropriate statistical analysis tools. Together with Statistics Sweden, Eurostat and universities we are looking into this research.

5. More specialized statistical studies for central and local government

19. In recent years, the number of requests for customized studies and the demand for microdata have grown sharply. To be able to fulfill the growing demand in both areas, the capacity of and services provided by our Centre for Policy-related statistics will be extended. The costs involved will be recovered in the fees charged for the work.

6. More timely statistics

20. Further improvement in timeliness remains an important goal. The use of new sources, big data and new ways of surveying are points of departure in this respect. In the process organization, too, reduction of throughput time and thus more timely results have a high priority. And again: users' wishes will be set against costs and potential effects on administrative burden.

D. Coherence

21. The increasing complexity of modern society has led to more demand for "evidence" on which policy making may be based; cf. Keuning and Zeelenberg, 2002. This inherent complexity leads also to a demand for coherent evidence. It is no longer sufficient to have data on several unconnected phenomena. Also, developments in society are much faster than before, and information travels very much faster than some decades ago. So we have user needs for more data that must be at the same time more coherent and timelier and must also show more of the dynamics of modern society.

22. Statistics Netherlands responds to this trend by shifting focus from individual statistics to coherent statistical programs describing the whole of a phenomenon in society. This approach will also identify gaps in the statistical program around this phenomenon as well as possible linkages between programs.

23. As an example, take communication and statistics about unemployment. Communication around unemployment focuses now more on the phenomenon of the labor market as a whole and less on unemployment alone. In the unemployment releases we also present data on the number of jobs, and we try to relate changes in unemployment to changes in the number of jobs. In the process of doing this, we became aware that what we really needed was data on the flows on, from and to the labor market, so that in effect we would have a complete picture of stocks and flows of the labor market and associated life-cycle statuses such as retirement, schooling, disguised unemployment, household work and so on. In technical statistical terms: we need an accounting system of the stocks and flows of the population according to the various activities of individuals. So we are now working on trying to complete the picture of the labor market, by integrating various research and development efforts.

E. Innovation

24. An important aspect of creating new statistics and production processes for them is the time needed for implementation. To answer quickly to new user demands, it is necessary to have flexible production processes that can be quickly implemented or adapted to new demands. To speed up this innovation process, Statistics Netherlands has founded an *innovation lab*, a unit for innovation, separate from research on the one hand and production on the other hand (Verbruggen and Braaksma, 2014; Braaksma et al, 2012). Such an innovation lab makes it possible to transfer ideas and tools from outside to inside as well as assist in implementing these in production and act as intermediary between research and implementation.

III. Conclusions and recommendations

25. Looking at the questions posed to the seminar, we may draw conclusions from the previous discussions.

A. How should statistical products evolve to respond to new demands, including information needs for policy decisions?

26. In the medium term, statistical product innovation should proceed through elaborate discussions with users. We should use not only existing channels, but also modern ones such as internet, focus groups, public consultation, and perhaps also social media such as Facebook.

27. In the short term we should try to keep up to date by intensive contacts with the main users, in particular ministries, and with the press.

B. What new data-based services need to be developed to enhance users ability to extract knowledge out of data?

28. We believe that open data will be the most important new service that enables users to extract data and use this in their own databases and applications. Besides this, better access to microdata, either onsite at the national statistical office or through remote access, will remain an important service, and continuous improvement is needed. Also, more advanced means of cooperation such as through secure multi-party communication needs more attention and research into its possibilities.

C. How can national statistical offices increase access, usability and re-use of data to strengthen value and quality?

29. There are two means to increase access and usability. The first is to adapt to new ways of communication; at present this means open data and smart phones, but we must keep up with future developments and identify new possibilities. This is an area where international cooperation can be especially fruitful.

30. Secondly, coherence between statistics and integrated description of social phenomena should be a high priority for national statistical offices. It is no longer sufficient to present statistics as such. On the contrary, we should describe a phenomenon with the help of coherent statistics, and also identify gaps where new statistics are needed to describe the phenomenon in question.

D. What are the key constraints (including legal) to promoting greater data integration to generate value for users?

31. The greatest challenge will be to maintain impartiality and equal access to all users, in particular when cooperating with private enterprises and public institutions.

E. Which could be the strategic partnerships with private sector to enhance tools for data-analytics as a service?

32. The most important strategic partnerships would be with the press, in particular in making special reports, forecasting, research and analysis agencies within or linked to the government, and private data-processing and information enterprises, such as Google, banks, and credit card companies.

33. For this cooperation to be successful, it is important that NSIs have a unit for innovation, separate from research on the one hand and production on the other hand. Such an "innovation lab" makes it possible to transfer ideas and tools from outside to inside as well as assist in implementing these in production and act as intermediary between research and implementation.

References

Arjan de Jong, Erik Schepens, Jos Schiepers, Jeroen van Velzen, Henk Verduin, Marga Verstege-Jilesen and Kees Zeelenberg, 2011, *Review of statistical priorities at Statistics Netherlands*. Discussion Paper 2011-07, Statistics Netherlands. www.cbs.nl/nl-NL/menu/methoden/onderzoek-methoden/discussionpapers/archief/2011/2011-x10-07-art.htm

Barteld Braaksma, Nico Heerschap, Marco Roos and Marleen Verbruggen, 2013, *Innovation at Statistics Netherlands*. Paper presented at the UN-ECE *Seminar on New Frontiers for Statistical Data Collection*. www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.44/2012/mtg2/WP7.pdf

Steven Keuning and Kees Zeelenberg, 2003, *Samenhang in de statistiek* (Coherence in statistics). In: A.P. Ros and H.R.J. Vollebergh (eds), *Ontwikkeling en Overheid*, SDU: The Hague, pp. 139-150.

Statistics Netherlands, 2013, *Strategic multi-annual programme 2014-2018*. www.cbs.nl/en-GB/menu/organisatie/corporate-informatie/publicaties/archief/2013/strategic-multi-annual-programme-2014-2018.htm

Marleen Verbuggen and Barteld Braaksma, *Innovation at Statistics Netherlands*. Invited presentation at the IAOS 2014 Conference Meeting the demands of a changing world, 8-10 October 2014, Da Nang, Viet Nam. https://iaos2014.gso.gov.vn/document/marl1.s1.v1.pptx

Wikipedia, 2015, *Secure multi-party computation*. http://en.wikipedia.org/wiki/Secure_multi-party_computation

United Nations ECE/CES/2015/42

Economic and Social Council

Distr.: General
7 May 2015

English only

Economic Commission for Europe

Conference of European Statisticians

Sixty-third plenary session
Geneva, 15-17 June 2015
Item 4 of the provisional agenda
Modernisation of statistical production and services and managing for efficiency

Innovative products and services for European statistics: challenges and opportunities

Note by Eurostat

Summary

This paper focuses on innovative products and services that are a key aspect of the modernisation of European statistics. Accelerating demand for information by different stakeholders is putting pressures on statistical organizations around the world. They need to deliver timelier, high-quality information that responds to varying user needs and supports evidence-based decision making. Consumers are increasingly empowered by advanced technologies that leverage the value added of data integration and analytics for knowledge generation. To respond to this fast changing demand from users, statistical organizations need to strengthen the portfolio of products they offer by adding value to the statistics they disseminate.

This paper is presented for discussion to the second session of the Conference of European Statisticians' seminar "Modernisation of statistical production and services and managing for efficiency".

GE.15-09142 (E)

1509142

Please recycle

I. Introduction

1. In a digital and information driven society, official statistics face many challenges in responding to user's needs relating to technology, methodology, governance, organization of work, dissemination and innovation. Currently, national statistical offices and international statistical organizations are striving to adapt to the evolving society to stay relevant by becoming more modern.

2. The provision of innovative products and services is a key aspect of the modernisation of European statistics. The key drivers are: a progressive switch from statistics to dashboards of indicators produced by combining different sources and methods in a flexible and agile way; a renewed emphasis on timeliness combined with quality assurance; a value added approach to statistics moving towards the provision of statistical information and services; an interconnected view that encompasses basic statistics, accounting frameworks and indicators; products that are not anymore mere numbers but combine also metadata and visualization.

3. The European Statistical System (ESS) is moving along these new societal trends: until recently rapid estimates, composite indicators, scoreboards, wiki articles and mobile services represented the innovative strand in official statistics. This paper explores some potentially new ideas for renewed products and services of European statistics: portals, interlinked indicators, partnership-based outputs, classic statistics produced with innovative techniques, infographics, tailor-made statistics, linked open data (LOD) based dissemination and data analytics services.

4. For official statisticians, this is a new way of conceiving, producing and disseminating statistical information that reflects the technological and communication-oriented daily reality of societies.

5. These technology trends, coupled with the increased demand from users, are a great opportunity for statistical organizations to enhance their relevance as data are becoming a key source of knowledge, the new capital. Statistical offices will succeed in achieving this goal only if they are available to invest in strengthening their product and service portfolio, putting users in the driver seat and in recognising that quality is the overarching principle to ensure that products and services are valuable to society.

II. Innovative products and services

6. In the aftermath of the financial and economic crisis, official statisticians underwent an in-depth reflection on how to respond in an agile, flexible and tailored way to the increasing request for statistical information. Such an analysis targeted different dimensions of collection, compilation and dissemination of official statistics with a two-fold aim: First, to streamline the production of official statistics by optimising the production processes and, second, enhance the capacity of national and international statistical organizations in disseminating and communicating statistical products and services. The digital transformation of the society, enabled by the data revolution, is highly influencing these developments.

7. The following section will address the common features that can be identified in the different modernisation programmes that statistical offices have launched in the recent years, both as concerns statistical production and statistical output (products and services).

A. Recent trends

8. As regards the dissemination and communication of the output of statistical production, general trends could be highlighted.

1. Statistical data for policy making: indicators vs. dashboards vs. synthetic/composite indicators

9. The demand of indicators for policy making purposes has considerably increased during the last years. Evidence-based policy making has become a "must" in running economic, social and environmental policies at global, international, national and regional level. In a user-producer-dialogue, policies are more and more defined in "measurable" terms and associated to a set of indicators selected or created to provide the necessary information on the achievement of the policy targets. In addition, the achievement of policy targets is often based on thresholds derived from a statistical analysis and assessment of past performance.

10. If until few years ago, single key indicators were traditionally used for these purposes, nowadays the following types of indicator groups are developed to ensure a proper follow-up and measurement of efficacy and success of policies:

- Scoreboards (set of indicators on the situation);
- Dashboards (set of indicators for monitoring and control);
- Indicators catering for sub-populations (to measure distributive impacts – e.g. inflation by household spending group to measure distributive impacts of inflation);
- Composite indicators (compilation of individual indicators into a single index, on the basis of an underlying model) and;
- Synthetic indicators (to condense the measurement of several sub-indicators into one value).

11. This tendency appears quite strongly in the instruments used by the European Commission to run and monitor the European policies. For example the European Semester relies on a scoreboard of 11 headline indicators measuring internal and external macroeconomic imbalances and competitiveness of EU countries. It is monitored with 28 auxiliary indicators. The Europe 2020 strategy, on the other hand, aims at achieving 5 high level targets and measures this on the basis of 9 headline indicators translated into national targets and indicators. Similarly, other key European policies such as the Single Market Monitoring, the Digital Agenda, etc. are measured by statistical indicators. For all these instruments, Eurostat provides the corresponding dashboard and scoreboard indicators in a user friendly and policy oriented fashion, combining data and metadata and offering, in specific cases, quality indications on the indicators themselves.

12. Whilst traditionally indicator sets are quite well established in statistical offices, with successful examples such as those published by Eurostat (Principal European Economic Indicators, Europe 2020 indicators, GDP and beyond indicators, Macroeconomic Imbalance Procedure indicators), users have recently started to value the following new characteristics:

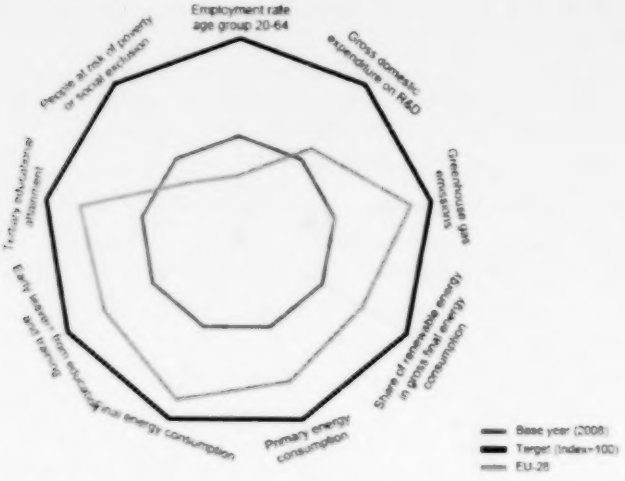
- Inclusiveness;
- Interlinks among different dimensions of the policy to be monitored;
- Agility and flexibility to respond to changing requirements;
- Attractiveness for policy makers (quality, including timeliness and accuracy; possibility to derive new indicators in a quicker and targeted manner);

- Easy accessibility;
- Synthesis.

13. These new characteristics call for agile production systems based on a smart and optimised use of the available statistical information and able to amplify the synergies in deriving "new" indicators from the existing set of basic statistics.

14. The challenge for statistical offices in this context is to set up production and information systems that will shift from the traditional paradigm of offering "statistics" to a more service oriented attitude "to connect, aggregate and tailor" statistical information, on the basis of a data warehouse approach and providing services in addition to data.

Figure 1
Examples of indicators disseminated by Eurostat


Employment rate age group 20-64
Gross domestic expenditure on R&D
Greenhouse gas emissions
Share of renewable energy in gross final energy consumption
Primary energy consumption
Final energy consumption
Early leavers from education and training
Tertiary educational attainment
People at risk of poverty or social exclusion
Base year (2008)
Target (Index=100)
EU-28

Europe 2020 indicators

GDP and Beyond indicators

Source: Eurostat

2. Rapid estimates

15. Timeliness has a prominent role among the quality dimensions associated to statistics and statistical indicators. The general perception about the timeliness aspect in the release of official statistics is quickly evolving in line with the trend of a more data driven society that relies on access to "numbers" on the fly. Indeed, if 25 years ago the switch, for example, from annual to quarterly national accounts took some time to be fully endorsed by users and producers, today fresh and quick statistical information is continuously required.

16. At the same time, new actors, mainly from the private sector, are starting to use new methodologies, techniques and data sources to produce proxies of official statistics. Concepts such as nowcasting and flash estimates are regularly used in statistical offices, and flash estimates of GDP and the Harmonised Index of Consumer Prices (HICP) are released by Eurostat for the European Union (EU) and euro area together with the available figures from EU member states. Rapid estimates and nowcasting are particularly important in the current economic and social conditions characterised by high volatility since they can provide fast feedback to policy makers on the policy impacts ("did we get it right or not?").

17. The challenge for the official statistics community is to ensure an adequate level of quality of the statistical indicators provided in a timely manner, as the quality may have consequences on the public image of and trust to statistical offices (a reflection linked to the idea of "labelling" of official statistics).

3. Modern dissemination and communication tools

18. Recently, particular emphasis has been put in statistical offices to adapting the style and services associated to dissemination and communication. The technological drivers, as well as the growing information demand, forced statistical offices to redesign the way of presenting and disseminating official statistics. Some offices have used professional communicators to focus their message and to help them adopt a modern style in communication. A number of statistical offices are showcasing new innovations: on one side by using new technological tools (interactive tools, mobile services, visualisation tools, etc.), and on the other side by applying a more targeted and modern communication approach (wiki-style dissemination sites, shorter and condensed presentations, almost disappearance of traditional publications, more visual-oriented presentation than traditional statistical tables, dynamic dissemination and data visualisation).

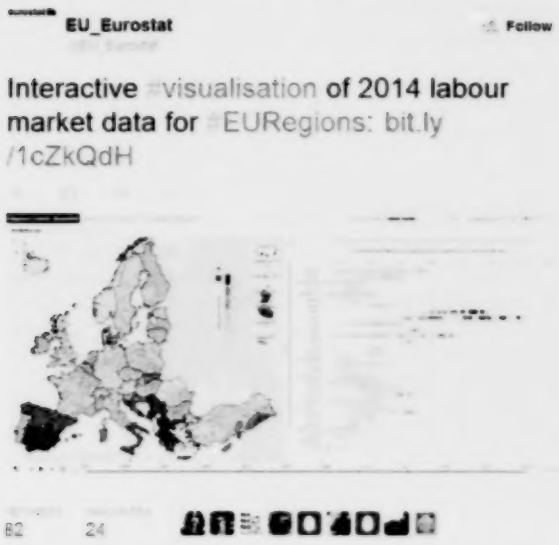
19. At the same time, the use of social networks for statistical dissemination (e.g. Twitter and Facebook) has brought a new dimension to disseminating statistics and in informing the general public with statistical news. Eventually, new profiles of data users appear, like the emergence of users of statistical data that create stories on the basis of data analytics services offered by statistical organizations (e.g. data journalists).

20. Eurostat is quite active in this context, for example

- Economic Trends offers a dynamic representation of a user-driven choice of the combination key indicator-country through a map visualisation;
- Statistics Illustrated provides a dynamic-over-the-time representation of key indicators (e.g. the MIP indicators), combining the evolution of the indicators by selected countries;
- Widgets give a quick information on key indicators (with the possibility of importing in its own web representation the widget facility);
- Inflation Dashboard provides a tool to display the Harmonised Consumer Price Index data simultaneously, by products and countries;

- Statistical Atlas offers an interactive map viewer for statistical and topographical maps;
- Mobile apps on the EU Economy and on Countries profile provides insights in European statistics through an easy access via mobile devices;
- Extraction tools - Tables, Graphs and Maps (TGM) offer a single style tool for downloads and prints of tables, graphs and maps;
- Data visualisation via Twitter addresses users via the social network;
- Infographics translate statistical information on key news and indicators in a visually oriented style.

Figure 2
Example of data visualisation tools provided by Eurostat



Source: Twitter.com/EU_Eurostat

Source: Eurostat Statistical Atlas

B. Future perspectives

21. Until now, official statisticians have quite successfully managed to innovate and develop new statistical products and services to meet new users' requirements. Nevertheless, the digital transformation and rapid changes that the society is experiencing require additional efforts. The following paragraphs briefly explore some potentially interesting ideas for the near future.

1. Portals

22. One of the key changes in the way of approaching statistical dissemination and communication in statistical offices is to offer tailor made services that generate added value to the traditional offer of statistics. In some countries, a tailor-made offer of statistical data and services targeting specific user groups, such as enterprises, has already started. The key characteristic of these services is to provide an attractive mix of dissemination facilities through a portal tailored to match the user interest: in case of enterprises, portals can offer specific tailor-made statistics for the interests of an enterprise or enterprises. The concept could be expanded to provide tailor-made services based on a data warehouse system that, in the ESS perspective, covers a system of statistical information repositories (data warehouses), is metadata driven and as much automated as possible. The Eurostat powered EU Census hub could be considered as a first realisation of such concept where users can define their queries and get comparable European information on census related indicators. Another example in this direction is the initiative "confidentiality on the fly" on which some statistical offices are currently working.

23. Most recently, portals increasingly offer diversified entry points for consumers and provide a guided tour that assists them in using the tool. In addition, portals provide a good data search and browsing capacity and are increasingly becoming search engines for statistics as opposed to a main websites of statistical offices.

24. The challenge for statistical offices is to set up the architecture and infrastructure required to successfully run portals that are user-friendly and serve a variety of data needs, as well as to switch from the traditional dissemination approach to a more tailored one.

2. Data integration

25. One of the major interests for users is getting information about the level of data integration. This information can be derived through metadata in single dissemination data warehouses. This approach could also serve the purpose of interlinking indicators, as a by-product.

26. Links between indicators describing different aspects of the same phenomenon are of increasing interest for users. This is particularly the case when analysing the impact of policies using a set of indicators.

27. The traditional narrow stovepipe approach usually applied by statistical offices in releasing statistical indicators has proven limited against the need for data integration. Measures of data integration and coherence are becoming a prerequisite for analysts. Therefore, statistical offices need to explore methodologies for better data integration to ensure, through the modernisation of statistical production, more integrated statistics (once again, a data warehouse organisation could facilitate this approach). In case of indicators, further steps along these lines are taken as the use of composite and synthetic indicators is becoming more widespread and benefits from interesting suggestions coming from other scientific disciplines (e.g. the use of POSET – partially ordered sets).

28. The challenge for statistical offices is to foster the development of methodological approaches to compile and describe sets of cross-domain indicators and to adequately represent and explain interlinks (in case on the basis of social and economic theories).

3. Partnership-based output

29. Statistics are not anymore exclusively produced by statistical authorities. The strong demand generated by a data-driven society attracts new major actors in the statistical market. Whilst we, official statisticians, could claim that the role of official statistics will continue to be relevant, notably as reference benchmark, we have to recognise that potential partnerships in collecting, processing and disseminating statistical products and services is an opportunity to be explored. Such partnerships could be useful for example in order to:

- Exploit the data revolution (in collaboration with big data collectors such as mobile phone providers or retail sale owners as regards their datasets);
- Benefit from powerful data dissemination partners (e.g. major web search engines or global vendors and their "data analytics" initiatives);
- Use privately produced information to derive statistics (e.g. real estate agencies data for compiling commercial property price indices).

30. For this purpose, Eurostat has successfully developed partnerships with, for example, a major web research engine for the dissemination of 10 sets of indicators (Google Public Data Explorer).

31. The challenges for statistical offices in this context vary according to the collaborative mechanisms that could be set up. Aspects like legal agreements, reputational effects and effective partnership have to be considered in this respect.

4. New data sources as generators of innovative products

32. New products and services could be developed using the available new sources, such as big data, administrative sources and geo-spatial information. New data sources can help generate new information: for example higher frequency information using big data (consumer sentiment on a weekly basis) or more geographical detail (territorial detailed labour force information for local policy making using administrative data combined with the labour force survey).

33. Eurostat is, with the ESS partners, currently exploring the potential offered by:

- New sources in particular by big data (e.g. tourism statistics based on mobile phone data and forecasts of key macroeconomic indicators using big data);
- The more intensive use of administrative data;
- Geo-referencing statistical data (e.g. the Inspire geo-portal, a EU Commission initiative).

34. The challenges for official statisticians in this context are mainly related to the resources to be invested in the methodology, technology and infrastructure as well as the required skills for the future generation of statisticians (data scientists).

5. Infographics, visualisation tools, media supports

35. Visual communication has become paramount in modern statistical communication and dissemination. Newspapers, media and presentations more and more make use of visual concepts, videos and new tools like infographics (visual representation of information). The opportunities offered by visual communication reflect the media-oriented attitude of the society as well as the intensive use of mobile applications and social networks. Whilst some statistical offices have started to disseminate official statistics using these tools, work still has to be done to keep the pace with the developments in the media.

36. In addition, new forms of communication, ranging from e-books to interactive press releases, focus on faster and more immediate communication between data and text, and build on data integration moving away from stovepipe or a topic by topic approach to dissemination. Social media press releases are a good example of this trend.

37. Eurosat is a front liner in exploring and deploying new visualisation tools that address these challenges: infographics for dedicated dissemination; videos to illustrate concepts and data; statistics explained (wiki dissemination of statistical information).

38. The challenge for statistical offices in this area is to be able to translate complex statistical concepts in attractive statistical information, on the basis of an appropriate choice of subjects, conception of synthesised messages combined with attractive representation and dissemination facilities.

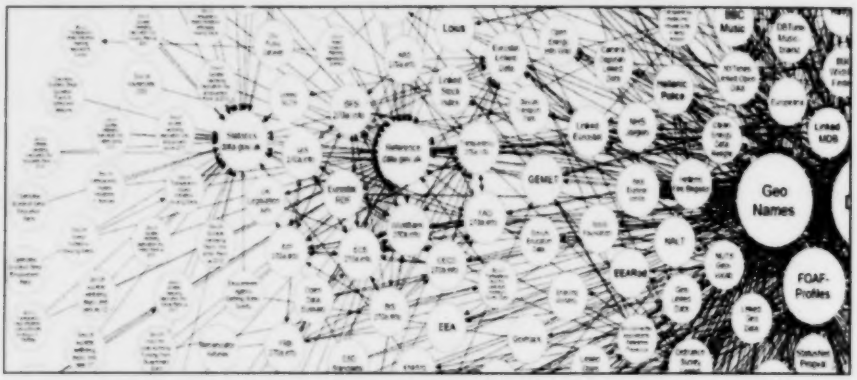
6. Linked Open Data (LOD) - semantic web

39. Linked Data is about using the Web to connect related data that wasn't previously linked, or using the Web to lower the barriers to linking data currently linked using other methods. More specifically, Linked Data is defined in Wikipedia as "a recommended best practice for exposing, sharing, and connecting pieces of data, information, and knowledge on the Semantic Web" using data and metadata standards.

40. From a statistical point of view the Linked Open Data (LOD) initiative builds on the general tendency of governmental organisations and statistical offices to provide open data (for example, Eurostat data represents 90% of the EU open data portal offer) and to apply to them a web semantic approach. The semantic web is an extension of the current web in which information is given well-defined meaning, better enabling computers and people to work in cooperation. Such an approach would foster the link among data sets and in the end, would create a larger statistical information-based virtual data storage that could be exploited for statistical purposes, also according to a machine learning approach. In addition, LOD would favour the potential to generate reuse of data by semantic integration, catering to different users also through secondary dissemination platforms.

41. The challenge for statistical offices will be mainly in running the standardisation wave and couple it with the web opportunities.

Figure 3
Linked Open Data



Extract from the Linking Open Data cloud diagram 2014, by Max Schmachtenberg, Christian Bizer, Anja Jentzsch and Richard Cyganiak. http://lod-cloud.net/

7. Machine learning

42. Technical advancement, such as cloud services and computing, could facilitate a statistical use of data available on the web. Recently, a big web retailer (Amazon) launched a new cloud service to help users build predictions on the basis of model tools that use data mining techniques. Users can use the service for a price and harness their own data sources. This is related to the move towards a decision system support in a cloud environment and it is close to the concept of data analytics as a service that is becoming very relevant for official statistics.

43. The challenge for official statisticians is to create a cloud platform where data are stored and analysed in an integrated way via data mining models that offer description and prediction tailored to users (e.g. client profiling, behaviour responses). A further step towards a service oriented approach for statistical offices.

8. Data analytics services

44. Statistical offices are becoming statistical service providers. "Statistics as a service" will represent an increasingly important development along with cloud environments for data integration and data analytics. The "platform as a service" type of framework will be increasingly introduced in the statistical community, in line with the expectation that processing efficiency will increase more rapidly than data transfer capacity (e.g. data gravity models). This leads to the creation of large aggregate data centres on cloud platforms. Data analytics combined with web semantics represents the future research development in this area.

45. The challenge for official statisticians is to address these developments from a statistical point of view, with consideration to privacy, confidentiality, legal aspects, analytical developments, provision and use by data owners and data analytics platforms for analysis and prediction as well as building capacity for data analytics.

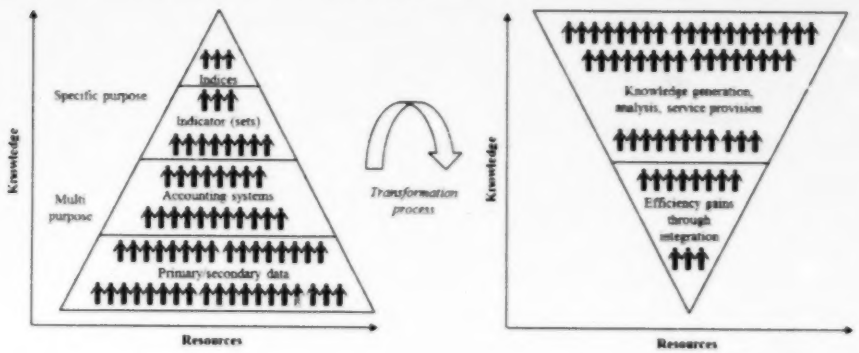
III. Issues and challenges

46. The introduction of innovative products and services in the official statistics community generates specific challenges. The main challenge is creating working environments that are oriented to creativity and innovation, support the development of new products and enhance the skills of staff to perform tasks from data science to communication in addition to new and consolidated methodological skills.

47. Another key challenge is to make the statistical organisation evolve and successfully undertake a paradigm shift from the traditional way of producing statistics to a more service oriented culture based on modern methods and techniques. The challenge will be to move from the current system towards a knowledge-oriented statistical service, emphasising the orientation towards the clients, i.e. to serve the users in the best possible way. This implies a shift in the focus of statistical offices from an internal perspective (production system) to an external one (users). Organisationally speaking, this paradigm shift relies on strengthening the services layer through a gradual transformation process targeted to generate knowledge, analysis and services.

Figure 3

Transformation process in statistical organisations



IV. Conclusions

48. In order to address these challenges from a broader perspective to modernisation, ESS undertook a major initiative to meet users' requirements and face the methodological and technological challenges to develop its products and services towards the ESS Vision 2020. The ESS Vision 2020 is the framework for the modernisation of the ESS through five strategic goals relating to users, new sources, quality, production processes and dissemination/communication. This Vision is currently being implemented through a portfolio of enabling projects to build the ESS capabilities for achieving the Vision.

49. One of the key strands for modernisation in the ESS focuses on digital communication, users, analytics and innovative products. This programme will facilitate the move from a traditional mix of statistics to innovative products and services based on a user-driven, technologically advanced and service-oriented culture, supported by adequate infrastructure and architecture. This strand, combined with new data sources and quality framework, is planned to bring innovation into ESS products and services. The achievement of the ESS Vision 2020 depends on the ability of the ESS to be open to innovation and successfully apply a change management in a collaborative effort.

United Nations ECE/CES/2015/43

Economic and Social Council

Distr.: General
1 June 2015

English only

Economic Commission for Europe

Conference of European Statisticians

Sixty-third plenary session
Geneva, 15-17 June 2015
Item 4 of the provisional agenda
Modernization of statistical production and services and managing for efficiency

Annual report of the High-Level Group for the Modernization of Statistical Production and Services

Note by the High-Level Group

Summary

This paper informs the Conference of European Statisticians about the activities and outputs of the High-Level Group for the Modernization of Statistical Production and Services. The purpose of the Group is to guide strategic developments in the ways official statistics are produced.

The Conference of European Statisticians is invited to express its views on the value of the work done so far and provide guidance on possible future priorities.

GE.15-11047 (E)

1511047

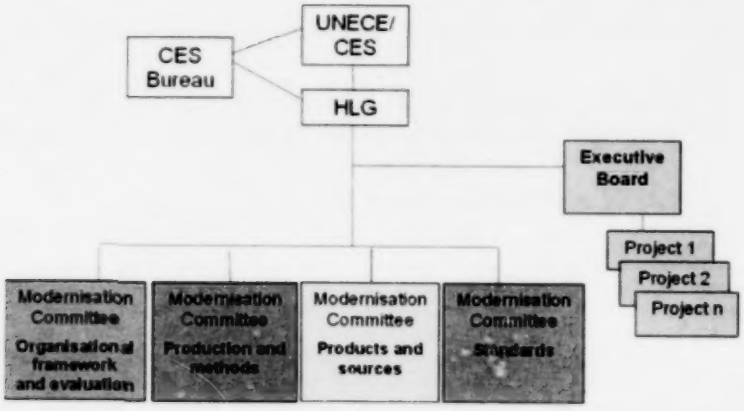
Please recycle

I. Introduction

1. The High-level Group for the Modernization of Statistical Production and Services (HLG) was created by the Bureau of the Conference of European Statisticians in 2010. It comprises the heads of eleven national and international statistical organizations, and has a mandate to reflect on and guide strategic developments in the ways in which official statistics are produced.

2. To further the work on statistical modernization, the HLG oversees four modernization committees, as well as annual multi-national collaboration projects, which are managed by an Executive Board. Figure 1 illustrates this governance structure, which was endorsed by the Conference of European Statisticians in 2013.

Figure 1
Governance structure for HLG activities



3. To determine the topics for the projects, the HLG organises a workshop each November, inviting representatives of various expert groups and projects related to modernization of official statistics. These workshops review progress and determine the key priorities for the following year. In November 2013, the workshop decided that implementing the Common Statistical Production Architecture (CSPA)[1] and investigating the potential use of Big Data sources for official statistics were the two highest priorities for 2014. In November 2014, the priorities decided for 2015 were further work on the implementation of the Common Statistical Production Architecture, and development of the Big Data "Sandbox" for the production of "real" statistics.

II. Common Statistical Production Architecture

4. The Common Statistical Production Architecture (CSPA) was created during 2013 in an international collaboration project under HLG. That project included the

[1] For more information about how the CSPA supports the modernization of official statistics by facilitating the concept of "plug and play" for the components of statistical production, please see http://www1.unece.org/stat/platform/display/CSPA

development and specification of the architecture, as well as a practical test of its principles and applicability in a "proof of concept". However, given the strict time limits for that project (one year), it deliberately did not include the implementation of the architecture for production systems within statistical organizations.

5. CSPA provides a standard framework for developing the components of statistical production in a way that they can be much more easily shared within and between organizations. It builds on key international standards such as the Generic Statistical Business Process Model (GSBPM), and the Generic Statistical Information Model (GSIM).

6. During the 2014 project the following CSPA-compliant services were developed:

- Seasonal adjustment – France, Australia, New Zealand
- Confidentialized analysis of microdata – Canada, Australia
- Statistical chart generator – OECD
- SDMX transformer – OECD
- Sample selection – Netherlands
- Linear error localisation – Netherlands
- Linear rule checking – Netherlands
- Error correction – Italy

7. These services have been incorporated into a "service catalogue", which is hosted by Eurostat, and are available for use by other statistical organizations. Another result of the 2014 project was a slightly revised version of CSPA (version 1.1), which incorporates additional guidance based on the lessons learned from the project.

8. The 2015 project has the following objectives:

- To extend the governance and support offered for the implementation of CSPA compliant statistical services;
- To build more CSPA services and have a plan in place for continued investment in the development of CSPA services;
- To facilitate the transitioning of CSPA governance to the Modernization Committee on Production and Methods.

9. To achieve these objectives, work is currently focusing on:

- Developing a common investment strategy for participating organizations based on greater alignment of business plans and identifying common capability requirements;
- Creating a CSPA Logical Information Model to facilitate service specification and implementation;
- Developing new CSPA services, e.g. for probabilistic record linkage, classification management, and validation rules specification.

III. Big Data

10. The HLG has identified the need to better understand the impact of new data sources, and in particular "Big Data" on official statistics. Following the

endorsement of the HLG paper "What does Big Data mean for official statistics?" (ECE/CES/2013/20) by the Conference of European Statisticians, a project was launched in 2014, with three main objectives:

(a) To identify, examine and provide guidance for statistical organizations on the main possibilities offered by Big Data and to act upon the main strategic and methodological issues that Big Data poses for the official statistics industry;

(b) To demonstrate the feasibility of efficient production of both novel products and 'mainstream' official statistics using Big Data sources, and the possibility to replicate these approaches across different national contexts;

(c) To facilitate the sharing across organizations of knowledge, expertise, tools and methods for the production of statistics using Big Data sources.

11. The main outcomes of the 2014 project were:

(a) A framework for assessing the quality of Big Data;

(b) Papers on the current status of statistical disclosure control, the characteristics of Big Data and their implications for data privacy, and practical measures to manage Big Data privacy;

(c) Guidelines for the establishment of partnerships in Big Data projects for official statistics;

(d) A survey on skills and training needed to work with Big Data, and a draft skills profile for data scientists;

(e) The "Sandbox", a state of the art web-accessible computer environment, created in partnership with the Irish Central Statistics Office and the Irish Centre for High-End Computing, which provides a platform for collaboration where participants can experiment with different types of Big Data, and learn how to use the latest software tools for Big Data processing.

12. All results of the 2014 project, including the outcomes of the initial Sandbox experiments are available at:

http://www1.unece.org/stat/platform/display/bigdata/2014+Project

13. The 2015 project is focusing much more on the use of the Sandbox as an international resource for Big Data experiments and training, though it also has the very concrete goal of producing a set of internationally comparable statistics from one or more Big Data sources by the end of 2015. The scope of the project was deliberately defined to ensure coordination with related projects being led by Eurostat and the UN Statistical Division, who are both active partners in the HLG Big Data work. Current work includes investigation of data from Twitter and Wikipedia, as well as potential efficiency savings using Big Data tools to process traditional sources such as the UN Statistical Division foreign trade data.

14. The value of the Sandbox as a resource for the international statistical community is being increasingly recognised. A "sprint" session will be held in Cork, Ireland, on 22-24 June, bringing together representatives of key national and international statistical organizations, to develop proposals on the future governance and funding model for the Sandbox.

IV. The work of the modernization committees

A. Modernization Committee on Organization Framework and Evaluation

15. Key items of work completed or currently in progress within the modernization committee include:

- Survey on best practices in change and risk management;
- Study on legal and licensing issues, which led to proposals for a statement of intent;
- Generic skills profile for staff working with Big Data in statistical organizations is currently being tested;
- Guidelines for managers, including best practices, under development.

B. Modernization Committee on Production and Methods

16. Key items of work completed or currently in progress within the modernization committee include:

- Paper on the state of the art in the use of machine learning in statistical organizations;
- Generic Framework for Statistical Data Editing – draft prepared for peer review;
- Modernization maturity model - initial work to develop a model to assess the state of readiness of a statistical organization to adopt HLG outputs.

C. Modernization Committee on Products and Sources

17. Key items of work completed or currently in progress within the modernization committee include:

- Inventory of Big Data activities in statistical organizations;
- Study on the use of mobile devices for data collection and dissemination;
- Papers and survey of best practices on marketing and communicating the value of official statistics;
- Best practices in linked open data, including formalising and publishing key metadata in linked open data formats.

D. Modernization Committee on Standards

18. Key items of work completed or currently in progress within the modernization committee include:

- Generic Activity Model for Statistical Organizations (GAMSO) – Version 1.0 released in March 2015;
- Glossary of metadata / modernization terminology - under development;

- Quality indicators for each sub-process of the Generic Statistical Business Process Model – first version released for public consultation.

19. In addition to the four modernization committees, the HLG has established a cross-cutting task team on communicating modernization, which is looking at how to improve the flow of knowledge about modernization activities within the official statistics community. It will report at the end of 2015.

V. Ensuring coordination

20. All of the activities listed above are carried out in close cooperation with other international organizations and initiatives. For example, Eurostat and OECD are members of the HLG. Eurostat staff are involved in almost all of the activities mentioned, whilst OECD, the European Central Bank, the UN Statistical Division, the World Bank, the World Trade Organization, and the UN Industrial Development Organization (UNIDO) are involved in those activities that are relevant to their work. There is close cooperation with the UN Economic and Social Commission for Asia and the Pacific, which has established mechanisms to spread information about modernization activities to its member countries.

21. These activities are not limited to those working in official statistics. Where relevant, partnerships are also forged with external bodies. For example the Data Documentation Initiative and academia are represented on the modernization committees, whilst the partnerships with the Irish Centre for High-End Computing, data providers and software companies have been essential to the success of the Big Data projects.

VI. Looking forwards

22. There are several recent developments that could influence the priorities for the HLG in 2016 and beyond:

(a) The "Statistical Modernization Community" – a proposal to put institutional participation in modernization activities on a more formal basis by introducing a "statement of intent" which organizations can sign up to. In many respects, this would simply formalise the current situation, with a clearer demonstration of support for the activities of the HLG;

(b) Developing the notion of "competencies" as the key "units" for investment planning and collaborative development. Competencies are defined as things that an organization does, or wants to do. They are combinations of "factors of production" such as methods, technology, people, skills and standards. An example of a statistical competency is seasonal adjustment. Joint work with the European Statistical System Task Force on Enterprise Architecture is foreseen to further elaborate a competency framework;

(c) Increased sharing and coordination of investment planning. If statistical organizations have easy access to information on who is investing, or planning to invest in improving capabilities in specific areas, formal and informal collaboration opportunities should be much easier to identify;

(d) Data integration. The work on Big Data and other new data sources (and modes of collection) shows that data integration from multiple sources and mixed collection modes will become increasingly important for statistical organizations over the next few years. Developing and enhancing data integration

techniques, methods and tools could be the subject of an international collaboration project in 2016;

(e) Reporting. The UN Statistical Commission decided in March 2015 to request regular reports on statistical modernization work from the Conference of European Statisticians. This will increase the global outreach of HLG activities.

VII. Input required from the Conference of European Statisticians

23. The Conference is invited to express its views on the value of the work done so far under the HLG and provide guidance on possible future priorities.

United Nations ECE/CES/2015/44

Economic and Social Council

Distr.: General
5 June 2015

English only

Economic Commission for Europe

Conference of European Statisticians

Sixty-third plenary session
Geneva, 15-17 June 2015
Item 3 of the provisional agenda
Response by official statistics to the Sustainable Development Goals

Sustainable Development Goals and data from administrative registers

Note by Statistics Denmark

Summary

This paper suggests that handling data challenges emerging from the Sustainable Development Goals (SDGs) could be approached by creating pools of data derived from existing statistical data sets and administrative registers. Data from these pools of data could be used as an input for the measurement of some of the Sustainable Development Goals indicators. Thinking along these lines may help when statisticians are faced with the challenge of measuring the Sustainable Development Goals and the related targets. More importantly, these data would help empower local administration and political decision making.

The paper is presented for discussion to the first session of the Conference of European Statisticians' seminar "Response by official statistics to the Sustainable Development Goals".

GE.15-09035 (E)
1509035

Please recycle

I. Introduction

1. Data sources and especially the potential of Big Data to impact the collection of data for SDGs has been an issue raised by several people and organizations. There are high hopes that Big Data will be able to solve at least some of the data challenges related to SDGs. Data challenges are a valid concern as the Millennium Development Goals (MDGs), which were less ambitious in scope, had major problems in data coverage. Much of the information covering MDGs is extrapolated; the Human Development Index is based on data three years old – and just as an illustration of the challenges facing us – about half of the children in sub-Saharan Africa are not registered at birth. In other words, there is a strong need to develop the ability to produce, coordinate and communicate data from official statistics around the world. SDGs are a major opportunity to address those needs.

2. This paper argues that one way of dealing with the challenges is to increase the capacity to produce official statistics based on data from administrative registers. To be able to do this, one needs both the registers as well as the capacity to use them as the basis for the production of statistics. On the one hand, the establishment of one or more core administrative registers will be a great advantage for countries without any register capacity. On the other hand, the ability to produce statistics based on administrative data can be learnt by all countries – as it is already the basis for statistical production especially in a number of developed countries.

3. Stiglitz is quoted for stating that: "*What we measure affects what we do. If we have wrong metrics, we will strive for the wrong things*" when he, together with others, presented the analysis and suggestions that led to a major focus among national statistical offices (NSOs), especially in the developed world, to measure well-being. This development is a recent example of how political focus on a specific area has led to the creation of new data and new political discussions. The development of the SDGs is a more ambitious example of an effort to get the right focus and later the right metrics. And getting the right metrics for SDGs is a challenge that we are all faced with.

4. Big Data have often been mentioned when discussing these challenges. Big Data are an interesting new area of data with major potential that we have only started to explore. Big Data are created, for example, when data are collected on the traffic on mobile devices like mobile phones, and in social media like Google, Facebook and Twitter. As we need to look for new ways to get data to cover the SDG indicators, Big Data are worth exploring. But no NSO today has any major experience with Big Data and very little if any official statistics is produced based on Big Data.

5. This is not the case when it comes to statistics based on data from administrative registers. A number of countries including Denmark are already producing some. In the case of Denmark a lot of official statistics are based on data from administrative registers. But administrative registers are not only useful for the production of statistics. In fact, statistics are only a side product of administrative registers. The main purpose of administrative registers should be to provide a systematic and well informed basis for local administration and local political decision making.

6. The technological development that created Big Data might ease the collection of data inputs to create administrative registers. Mobile phones are wide spread all over the world, also in the developing world. The technological infrastructure that has been created through mobile networks could and should be used to get the data needed for the administrative registers. In that we can combine a well-known technology -- the mobile phone network – with another the technology for the production of statistics based on using administrative registers.

7. First, this paper presents a fully-fledged model for the production of statistics based on data from administrative registers.

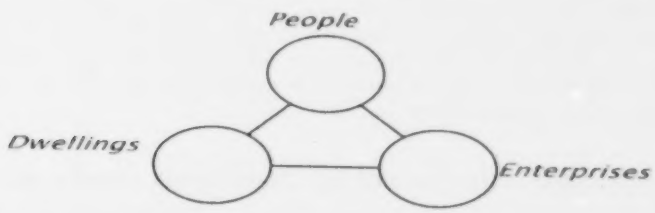
8. The paper then goes through some of the SDG indicators and argues that a population register or parts of a population register, such as a birth register, could lead to a marked improvement in the data coverage for the SDG indicators. And this is what we are striving for – to try to create a register-based statistical system that can, on the one hand, support local administration and, on the other hand, give us better data in relation to the SDG indicators. We want to – just like Stiglitz – do the right metrics.

II. Register-based statistics

9. Our point of departure is a fully fledged register-based statistical production system – but this is by no means our end goal. This short introduction to the full coverage system has two purposes. Firstly, to give the reader a sense of what is meant by a register-based statistical production system, and secondly, to relate the content of the full coverage system to the content covered by the SDGs in order to discuss how register-based statistics can contribute to the development of the statistical basis for SDGs.

10. The fully fledged register-based statistical production system is the one used today in Statistics Denmark. Close to 90 per cent of all social statistics and large parts of business statistics are produced based on data from administrative registers. The system is showed in figure 1.

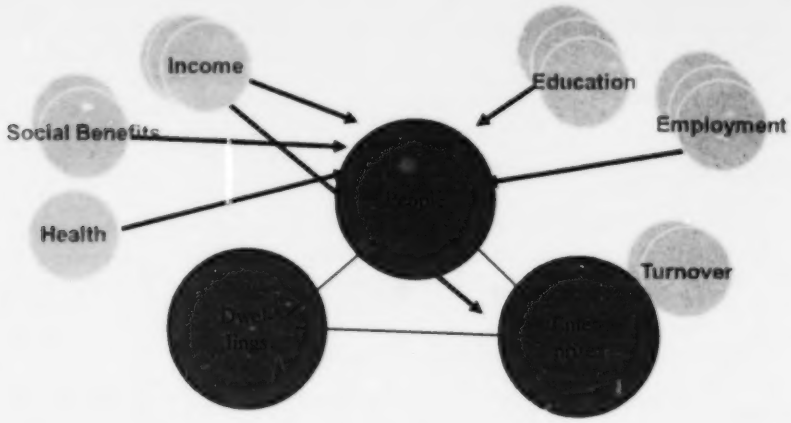
Figure 1
Core register structure in the Danish register-based production system



11. The core register structure of the Danish register-based production system contains three registers; a register on individuals (people), a register on active companies (enterprises) and a register on housing (dwellings). There is a unique identifier for each unit in the register (person, enterprise and building). This makes it possible to link information vertically between the registers, and horizontally within each register over time, to tell a story of the development of that specific unit in that specific register. At a given moment in time one can e.g. link a person to his or her dwelling as well as to his or her workplace (enterprise), and over time one can tell the statistical story of each building, enterprise or person from cradle to grave.

12. In Denmark it all began with the population register containing basic information on every individual, for example, place of birth, gender and biological parents. This basic demographic information can then be dynamically supplemented with information on e.g. marriage and divorce, children and migration. Each of the units in the three basic registers can be supplemented with a lot of other important information, of which most come from the population register i.e. information on individuals cf. figure 2.

Figure 2
Data related to the three core registers



13. The Danish register-based production system contains information on the health, social benefits both in cash and kind, income, educational background and employment situation of individuals – as well as information on the turnover of enterprises. This information can be combined in any given combination; the health and income situation of social benefit recipients; the educational background of people in full time or part time employment; the development in labor market participation for people with health problems; the development in enterprise turnover for enterprises, which increased the employment of workers with an academic background by 50 per cent. The possibilities are if not endless then at least numerous.

III. Register-based statistics and the Sustainable Development Goals

14. The system of goals and targets included in the SDGs is very ambitious. It is ambitious because it aims to cover close to all aspects of importance for human life, and because of the challenges it poses to NSOs worldwide and to their ability to provide data that can be used to monitor many target areas.

15. These challenges are an opportunity to come up with suggestions on how NSOs can handle the very ambitious agenda in relation to data that comes with SDGs. There is no simple solution to this, and it might be worthwhile to see the data challenges as an offset for a long term plan to develop statistical capabilities and competences worldwide. As is often stated, we tend to focus on the issues we measure and data are a very strong tool of measurement as well as a very strong means of communication.

16. This paper proposes to include data from administrative registers as one possible data source that could be developed to contribute to the data coverage of SDGs. This approach has several advantages and a number of challenges. The main advantage is that register-based statistics are cheap and flexible once the register has been established. Establishing the register is one of the challenges, and we will return to that. Another advantage is that the register will be able to support not only the production of statistics but also empower administration and policy development. The production of statistics based on

administrative registers is basically a byproduct of the establishment of the register, while the main purpose is to support administrative practices. A third advantage is that the running of a register is continuous. Data will be reported to the register continuously firstly to support the administrative processes and secondly to be used in statistics. For statistical purposes data are subject to quality assurance and documentation. The running of a well-functioning register ensures that data will not be 'three years old' as was mentioned earlier in relation to data used for the Human Development Index.

17. To establish a register is a challenge. It should be mentioned that whereas some countries are using data from administrative registers as an input to official statistics, there is a bigger group of countries were administrative registers already exist but are not necessarily used for statistics. They can be tax registers, registers on recipients of some kind of income support or local registers of university graduates. The existence of a register is of course a better starting point than if no registers exist at all.

18. There is no turn-key solution on how to establish a register. It could be based on methods used in countries that already have administrative registers. But it could also be based on a combination of existing knowledge and practices and new ways of doing things. If mobile data can help report 18 million births – as was the case in Nigeria in 2011-12 – this way of collecting digital data can be combined with the establishment of a register with continued reporting based on digital data collection through e.g. mobile devices.

19. As mentioned in the short introduction to the Danish register-based statistical production system, a core feature in a register-based production system is the population register. The establishment of a population register is in many ways a good place to begin the path to register-based statistics. This is also the case with the goals, targets and indicators in SDGs. In general, population is the denominator in many targets like e.g. 'mortality rate per 100,000', 'under-five mortality per 1,000 live births', 'percentage of women aged 20-24 who were married or in a union before age 18', 'percentage of population with electricity access', 'homicide and conflict-related deaths per 100,000 people'. A population register or just parts of a population register like a birth register will greatly help to reach the overall goal of empowering administration and facilitating data based political decisions. And on top of that it will be helpful in collecting data for the SDG targets.

20. There are other target areas in SDGs where one could consider establishing a register – again to support administration and political decision making but also to cater for the SDGs. A register of banks and ATMs as well as a population register could fulfill the data needs for the indicators 'proportion of population with an account at a formal financial institution' and 'number of commercial bank branches and ATMs per 100,000 adults'. A register of phone companies and their customers (the companies might already have that) is needed for the indicator 'individuals who own a mobile phone'. For indicators 'percentage of international cooperation projects being implemented to facilitate access to clean energy' and 'number of patents granted annually in developing countries...', a register on international projects in developing countries and, a register of patens would be useful. In other words, SDGs are a golden opportunity to develop registers to support the measurement and follow up of the SDG targets and indicators.

IV. Conclusions and recommendations

21. SDGs are putting data producers like the National Statistical Offices to a test. It will be a major challenge to be able to cover data needs of all the indicators involved in the SDG system. It is a challenge but it is also an opportunity. It is an opportunity for NSOs to work together in strengthening the capacity to cover the many different focus areas of

SDGs. This will not happen overnight. It calls for a mapping of what is already possible to cover, and it calls for initiatives on how to be able to cover what we are not able to cater for at present.

22. This paper suggests that some of that challenge should be covered by the creation of a number of data registers in core areas. Core areas not only for SDGs but also core areas that enable countries to administer and to take political decisions based on good data. There are a number of steps to this suggestion. There is a technical side as well as an administrative side to it. Technically the input interface for the register(s) has to be developed in accordance with the devices one wants to use to get data input to the register. Mobile phones and apps are useful in this sense, as mobile phones have spread all over the world, included the lesser developed parts of the world. Basically, the register should be set up to support daily practices at street level in all parts of health and social administration. These can be birth clinics, hospitals, local authorities administering social benefits in kind or cash or churches administering the registration of baptisms, marriages and funerals. This is the input side of the register, and this will only work as long as registration in the register helps these institutions in their daily work. The most important output of the register is data to support good administrative practices. Another output of the register is data that can be shared with the NSOs who can produce statistics based on the data from the register; statistics to support the measurement of SDG indicators.

23. Making this work calls for close collaboration between administrators at the local level, experts, who develop the technical part of the input to and the structure of the register as well as experts in the production of statistics based on data from administrative registers. As part of the activities needed to support the implementation of SDGs, a limited number of pilot projects along the lines described in this paper should be set up in collaboration with local and national administrators in a given country, and technical experts and experts from countries like Denmark, who already produce the majority of social statistics based on data from administrative registers.